2016

Modeling technologies and methods for DNA origami

Brian Thomas Nakayama
Iowa State University

Modeling technologies and methods for DNA origami

by

Brian Thomas Nakayama

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Computer Science

Program of Study Committee:

Robyn R. Lutz, Major Professor

Eric Henderson

Samik Basu

Iowa State University

Ames, Iowa

2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

Without the help of several people during my research, this thesis would not have been possible. In no particular order I'd like to thank the following people. I'd like to thank:

- Robyn Lutz for her wealth of knowledge and advice about graduate school and publication level writing, as well her many edits on this thesis and other papers not shown here.
- Divita Mathur for counting Origami structures and for imparting her knowledge and experience in wet lab techniques, along with her edits in appendix 2, her renderings of the Rectangle in fig. 3.9 and her help in implementing the experiment featured in and edits in chapter 3.
- Eric Henderson for counting Origami structures and for his advice and encouragement both in experiments and in academia in general.
- Drena Dobbs for her in depth reviews and discussion.
- Gabrielle Ortman for counting Origami structures.
- The anonymous reviewer from our PNAS submission for her/his helpful pointers.
- Other members of the Laboratory for Algorithmic Molecular Programming who have provided both discussion, and a space to practice presenting my work.
- Faculty of the Computer Science Department, who have done their best to make my experience at ISU a positive one while also helping me with the often confusing yet important paperwork.

This work was supported in part by NSF grants 1247051 and 1545028.

ABSTRACT

The creation of correctly assembling DNA origami often requires several iterations wherein a researcher tries and troubleshoots an incremental design. In each iteration there exists one or more costly failures that often take immense time or materials to find. These failures occur in part due to a lack of in-depth understanding of how DNA origami self-assembles and functions. To aid researchers in developing correct DNA origami designs, this thesis describes the creation of a DNA origami failure catalog as well as models for elucidating as-of-yet only partially understood properties of DNA origami. The failure catalog helps laboratory scientists gather requirements to preempt failures in their origami designs, and helps laboratory scientists troubleshoot their experiments after the implementation of a design by querying the catalog. Use of the catalog then helps verify the properties of new macro and micro models for DNA origami introduced here. These micro and macro models open up future ways to evaluate DNA origami through a mathematically more rigorous framework. By using both captured knowledge of previous design failures and novel theoretical modeling techniques, this work seeks to reduce the gap in understanding between design and implementation of DNA origami.

CHAPTER 1. INTRODUCTION

1.1 Overview

The design of DNA origami and structural DNA nanotechnology resembles closely the design of a computer program. Whereas a computer program encodes semantics using a 2 letter alphabet of 1's and 0's, DNA origami encodes shape and function (as in motion) using a 4 letter alphabet comprised of Adenosine, Thymine, Guanine and Cytosine as shown by Watson and Crick (1953). Just as a program can contain unexpected or unspecified deleterious behavior (Leveson (1995)), the process of self assembly of a DNA origami nanostructure can confound the nano-engineers who have carefully compiled a set of DNA sequences. As research in DNA origami increases, so too does the need for more comprehensive and sophisticated methods for determining the properties and failures of a given DNA origami design.

To this end, I have documented and formalized several unwanted structural properties in a DNA origami Failure Catalog. Chapter 2 details the creation and evaluation of the catalog. While the catalog exposes known errors and failures, it does not offer insight into yet unforeseen failures or properties of DNA origami. To do so one needs an expressive model that can capture subtle interactions and energy potentials that occur during folding.

Chapter 3 introduces a new model that I have developed with enough explanatory power to derive multi-state origami, where one designed DNA origami nanostructure using computer aided design (CAD, Douglas et al. (2009b)) software can result in multiple kinetically trapped native states. The model describes how to create design motifs for multiple native-state origami (multi-state) and identifies commonly used mono-state patterns. Chapter 3 also shows the results of a wetlab experiment similar to the one conducted by Endo et al. (2014) affirming the main ideas of our model. Chapter 4 summarizes the results and provides future directions.

1.2 Background

Research in molecular nanotechnology or nanomachines has followed the discovery (reviewed by Alberts (1998)) that a majority of living dry matter is composed of protein machines. Protein based enzymes (machines) rely on specific sequences of regular monomers such that under correct conditions (e.g. temperature) they self assemble or "fold" into a 3 dimensional shape with specific catalytic properties. These catalytic properties depend on the shape of the protein (Petsko and Ringe (2004)). Thus, in order to make new, protein-based nano-mechanical devices one must predict the structure created by an invented amino-acid sequence encoding a protein.

However, despite the abundance of protein based molecular motors found in nature, de novo techniques for constructing shapes from proteins have not progressed as rapidly as those inspired by DNA (King et al. (2012)). The relatively faster growth of DNA can be attributed to the highly reliable Waston-Crick base pairing interaction that leads to predictable secondary structure (Linko and Dietz (2013)), whereas predicting protein folding, i.e. the protein folding problem, has challenged researchers for decades (Levinthal (1968)). De novo DNA origami in particular has advanced to the point that researchers have made nanomechanical DNA devices with moving parts (reviewed by Kuzuya and Ohya (2014)).

DNA origami as introduced by Rothemund (2006) uses a 'one pot' reaction to fold a long single stranded DNA, called a *scaffold*, into a desired shape by binding with complementary short single strands of DNA, called *staples*. Originally, Rothemund used the 7,249 nucleotide genome of the m13mp18 virus as the scaffold; however, other researchers including Douglas et al. (2009a) have used longer genomes successfully. To design an origami object, first the researcher routes the scaffold into a desired shape. Often this step involves constructing an object from a set of straight parallel helices using software such as CaDNAno created by Douglas et al. (2009b). The scaffold is then routed through the helices. Places where the scaffold bends to connect one helix to another are called *crossovers*. Next staples are placed along the scaffold such that the backbone of the staples brings disparate domains, sequences of nucleotides, of the scaffold close together. Places where the backbone of a staple connects two disparate domains

are also called *crossovers*. Due to the periodicity of twist in β-form DNA, Rothemund (2006) assumes crossovers can only occur at certain positions to prevent prohibitive tension. The periodicity of twist will create over/under winding which generally constrains the design space by limiting the place where one can place crossovers. Under special circumstances one may purposefully create over/under wound structures (see Appendix A).

1.3 Tools and Models

Tools exist for designing shapes and generating staples (Douglas et al. (2009b) and Andersen et al. (2008)) as well as simulating the bending and movement of helices idealized as flexible rods (Kim et al. (2012) and Castro et al. (2011)). Recently, Yoo and Aksimentiev (2013) have also simulated DNA origami structures using molecular dynamics. Since the length of time for the formation of DNA origami takes minutes even under near perfect conditions (Sobczak et al. (2012)), creating molecular dynamics simulations capable of showing the entire folding process requires a highly coarse-grained model. Dunn et al. (2015) have modeled the folding of a very specific DNA origami using Continuous Time Markov Chains.

Several examples exist of DNA based nanostructures that move (Andersen et al. (2009), Kuzuya et al. (2011), Douglas et al. (2012), Zhou et al. (2015), Marras et al. (2015) and Saccà et al. (2015)) and leverage hydrophobic interactions to interact (Woo and Rothemund (2011) and Gerling et al. (2015)). Others have investigated ways to improve the *yield* of a specific DNA origami structure (Ke et al. (2012) and Sobczak et al. (2012)), the number of correctly formed structures. Research on improving yield along with research preventing errors from design to implementation of an origami nanostructure (Douglas et al. (2009b), Castro et al. (2011) and Doye et al. (2013)) supports the synthesis of functional DNA nanostructures. While the popularity and ease of creating nanostructures have led to many of the works above, the lessons learned from the production of DNA nanotechnology can carry over to the creation of nanotechnology using other molecules such as RNA (Geary and Andersen (2014)).

Models for DNA nanotechnology often rely on graphs wherein the nodes represent nucleotides and the edges represent various relationships between those nucleotides. Winfree (1998) used graphs to show the computational power of interacting single stranded DNA and

the DNA based Tile Assembly Model. Schaeffer (2012) used graphs as snapshots of a DNA system that folded according to a randomly sampled Continuous Time Markov Chain. For DNA origami, graph methods where nodes represent a crossover and edges represent double stranded helices of DNA have aided researchers inventing new structures (Zhang et al. (2015) and Benson et al. (2015)) and established the complexity of finding ideal scaffold routing schemes (Ellis-Monaghan et al. (2014)). The following chapter describes the creation of the failure catalog along with an in depth case study. The failure catalog influences the new model, as discussed in chapter 3.

CHAPTER 2. A FAILURE CATALOG FOR DNA ORIGAMI

2.1 Creating the Catalog

When designing a DNA origami experiment there exist many lessons learned from failures that have occurred previously and been published in literature. [1] By creating a Failure Catalog for DNA origami (Appendix A), new researchers entering the field can access relevant failures and methods for removing or mitigating a failure's effects. Below I describe the creation of a searchable failure catalog, and evaluation of a prototype called the "Slider" using the catalog.

The criteria for inclusion of a failure follows. Each failure in the catalog records (1) a significant failure, (2) described in the field's literature or laboratory notebooks, (3) that was caused in large part by a missing or incorrect goal or requirement, and (4) that has recurred, or could feasibly recur, in similar, future nanosystems. Many of the reported failures involve missing assumptions where domain knowledge did not exist at the time when the original failure occurred. Thus, inclusion of a failure in the catalog does not mean that the developer of the nanosystem in which the failure occurred was at fault, given the state of domain knowledge at the time. Rather, the effort is to disseminate the information about the problematic environmental assumption so that this failure will not recur.

Each failure diagram in the catalog contains the following fields:

- Assumed Properties
- Affected System Goals
- Failure Frame Type
- Description

[1] The following contains work done in collaboration with Thein Tun, Robyn Lutz, Yijun Yu, Divita Mathur and Bashar Nuseibeh (Tun et al. (2015)).

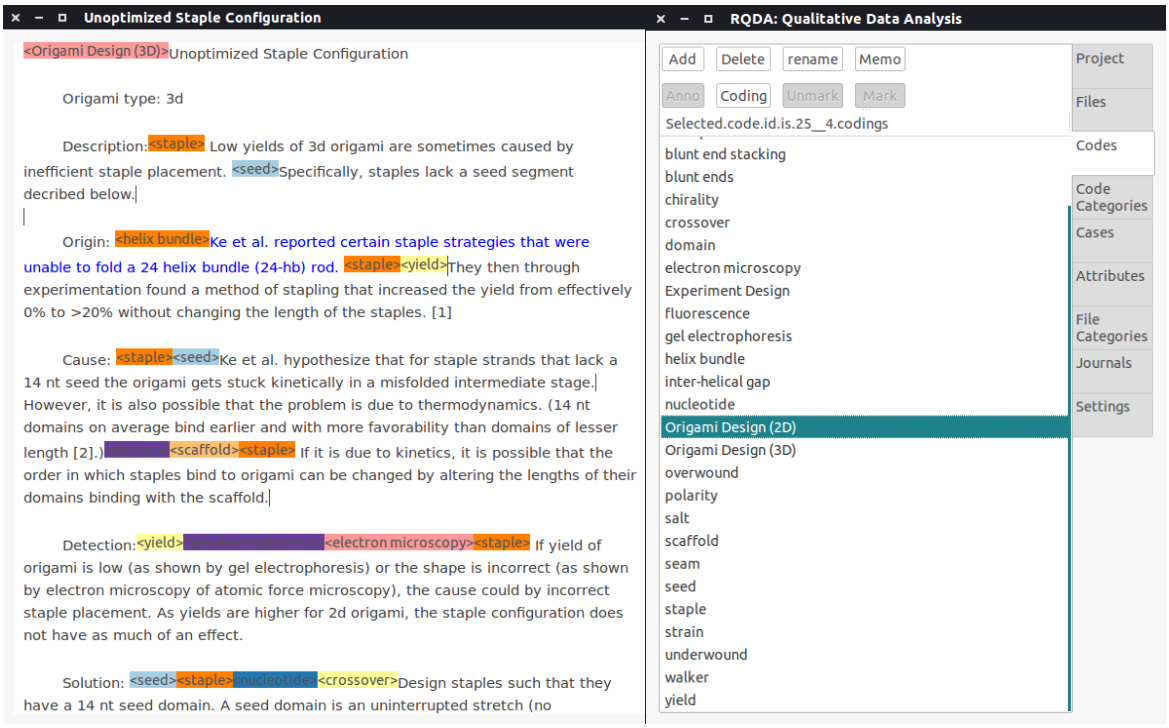


Figure 2.1 RQDA Catalog. *Left*: Example failure with "codes." *Right*: Example queries.

- Origin
- Cause
- Detection
- Solution/Mitigation

Assumed Properties lists the domain properties and assumptions that were missing or contributed to the failure. *Affected System Goals* contains the requirements whose satisfaction is obstructed by the failure. The *Failure Frame Type* contains the type of the failure defined later. The *Description* offers a short summary of the failure. *Origin* summarizes the literature while citing references for traceability. The *Cause* is similar to a failure mode in FMECA (Leveson (1995)), as it describes how the domains or their interaction led to the failure. *Detection* identifies how the effect of the failure was observed, and the *Solution/Mitigation* field contains the lessons learned that one can reuse to prevent or mitigate the failure.

I populated the catalog with lessons learned from journal articles and the articles' often lengthy (up to 100 pages or more) published supplementary information, as well as from lessons learned from a DNA origami researcher's laboratory experience in molecular biology. All failure entries also have traceability information to the published research describing the failed assumption and the proposed fix or mitigation strategy.

As defined in Tun et al. (2015), we further separate our failures into one of three failure types: *Component*, *Interaction*, and *Compound*. *Component* failures happen due to an incorrect specification, i.e., an observed, undesired property does not match a specified property. *Interaction* failures occur between the interfaces of two machines. *Compound* failures occur due to more than one incorrect specification, which combined lead to an undesired property. Each failure can also additionally be classified as *Cumulative*, a failure that occurs only if the frequency or accrual of an event happens above a specified threshold.

The catalog currently contains thirteen failures. Seven of these are of category (i.e., failure type) Component failures; three of these are of category Interaction failures; and three of these are of category Compound failures. Failures in any of the three categories can also be characterized as Cumulative failures if the failure only occurs beyond a certain threshold. Five of the failures are also of category Cumulative failures.

To store the catalog, we used RQDA (Huang (2012)). RQDA's coding function allows sources to be tagged with "codes" which one can later access through clicking. To make the catalog searchable, each coded page in RQDA represents one failure. For example, if one wanted to pull up all failures that affected the work of assembling DNA origami, one would click "Assembly Work" to access those failures.

The catalog is available online for use by the molecular programming and DNA nanotechnology communities (Nakayama (2015)). By making the repository publicly available, other domain experts can see failures, add new failures or refine current ones. The catalog has been reviewed by a domain expert as described below, but continued correctness of the repository requires ongoing review, and the catalog will change as our understanding of DNA based nanotechnology increases.

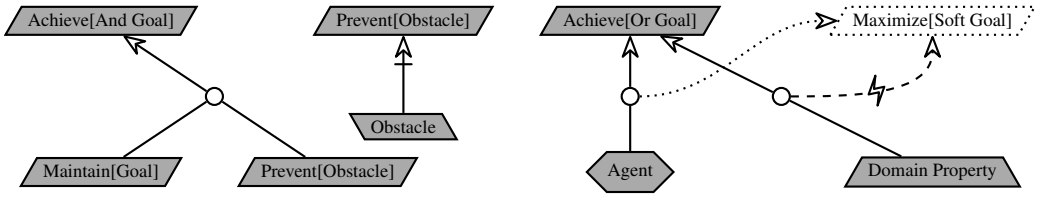


Figure 2.2 Goal Modeling Entities and Relationships

2.2 Evaluation

To evaluate the correctness of the catalog, a lead researcher of a DNA origami lab reviewed the entries in the catalog. All identified inaccuracies and missing information were corrected as per the expert's recommendations. We then evaluated the catalog retroactively on a real-world case study, a DNA origami device informally called the "Slider." Each of the *lessons learned* (LL) mapped one-to-one to one of the failures in the catalog. We represent the evolution of requirements using Lamsweerde (2009).

2.2.1 Goal Modeling

When modeling requirements as goals Lamsweerde (2009) often considers behavioral goals of three different variants: *Achieve, Maintain,* and *Prevent*, shown as parallelograms in Figure 2.2. A product must satisfy achieve goals by having the part in the brackets "[]" true at some time. Maintain goals are the same as achieve goals, except the product must *always* satisfy the part in brackets. A product satisfies a prevent goal by ensuring the part in brackets is always false. In addition to behavioral goals, we have two types of atomic properties: agents and domain properties/hypotheses, shown as trapezoids. An agent is any system component (such as a human or machine operator) that contributes to a goal. A domain property is an assumption that we expect to hold true, while a domain hypothesis is an assumption that is subject to change. Often when eliciting missing requirements, incorrect domain properties and hypotheses must be refined into finer grained statements.

The *And-Goal* in Figure 2.2 is *And-refined* into a maintain goal and a prevent goal. We always represent and-refinements with an edge from a circle pointing with an arrow to a *parent*

goal. *Child* goals have an edge without an arrow to the circle. Both child goals attached to the circle must be satisfied to satisfy the and-goal. In contrast, the *Or-Goal* is assigned to an agent and a domain property. The multiple arrows show the multiple ways the or-goal can be satisfied. Finally, the *Obstacle* prevents the satisfaction of a goal, i.e., if an obstacle is true, then the prevention of that obstacle will be false. Notice that the parallelograms used for obstacles have reversed angles. Furthermore, the arrow from the obstacle to the prevent goal has a bar going through it. This means the obstacle prevents the satisfaction of the parent goal.

When Or-refining a goal, each alternative satisfaction may add to or detract from a *Soft Goal*. Soft goals have a dotted border. Or-refinements that contribute to (help) a soft-goal have a dotted arrow from their circle to the soft-goal. Or-refinements that detract from (hurt) a soft-goal have a dashed arrow with a lightning bolt going through it. Soft goals are often used to choose between alternatives, or to differentiate between alternative products in a product line.

2.2.2 The Slider

The Slider moves a mobile platform composed of DNA upon interacting with a molecular input. The failures we examined focus on programming the self assembly of the slider from scaffold and staple strands.

The complex functional shape of the Slider has led to three main experimental failures. Each failure resulted in redesign of the device based on an incrementally updated goal diagram. We make the goal diagram explicit here in Figure 2.3 drawing inspiration from the goal models of Lamsweerde (2009).

Figure 2.3 shows the device after the first iteration. The first iteration of the Slider was a 2 dimensional piece of origami whose function required "inter-origami separation." That is, the design required each piece of origami to perform a task as a unit floating/existing separately in solution. Originally, the researcher involved assumed that the molecular properties of DNA fulfilled the implicit "inter-origami separation" goal as long as there did not exist a designed Watson-Crick pairing for inter-origami attachment. The researcher's internal working model of origami assumed there did not exist any other significant inter-nucleotide forces other than

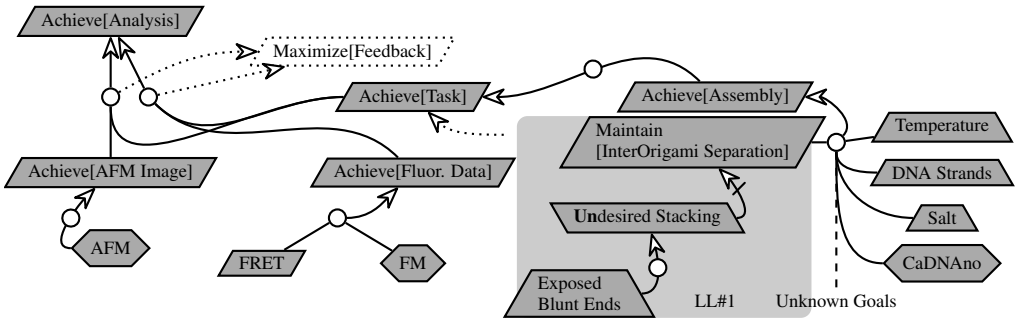


Figure 2.3 Goal Modeling for the Slider

Watson-Crick interactions. Further obscuring the requirement, the implicit assumption of inter-origami separation holds true in practice for other DNA-based nanotechnology experiments like tiling, as shown in Winfree (1998), or DNA strand displacement, as shown in Qian and Winfree (2011).

For the Slider, *stacking interactions* led to the obstruction of the "separation" goal (Failure VIII of Appendix A). DNA base-pair stacking in fact occurred due to understood hydrophobic/polar interactions between exposed adjacent base pairs on the edges of origami (Woo and Rothemund (2011)).

If the researcher had known of their stacking interaction via our catalog, they could have skipped the first iteration of the Slider. According to the researcher involved, two sources of information on stacking existed before the first iteration: the Rothemund (2006) paper on DNA origami and Endo et al. (2010). In Rothemund's original paper, the information about stacking was not the focus of the paper. The other merits of Rothemund's paper, such as the discovery of high yields and methodology for creating origami, meant that the stacking interaction was not emphasized or understood to the extent required by the Slider. Similarly, the origami design by Endo et al. noted the effect of base-pair stacking when their origami formed "extended structures". However, they did not explicate upon how one might fix base pair stacking.

Notably, a more thorough paper on stacking interactions, Woo and Rothemund (2011), was not yet available. As evidenced by Endo et al. (2010) researchers already knew about stacking, as they identify π stacking as the cause of extended (multi-origami) structures. Preventing the

experimental failure in this case could have saved the researcher working on the Slider $2,000 and 3 days. To correct the experimental failure, the blunt ends were eliminated (consistent with Failure VIII of Appendix A), preventing the fulfillment of the obstacle in the first iteration shown as LL#1 in Figure 2.3.

The second iteration of the slider produced unexpected Förster Resonance Energy Transfer values, thereby failing analysis. As the top level goal of conducting a DNA origami experiment, the achievement of correct analysis relies on the correctness of all contributing subgoals. It led to an inquiry of missing goals leading to the hypothesis that 2 dimensional origami had a prohibitively high amount of flexibility, limiting the predictability of the design. Before this hypothesis, researchers working on the Slider assumed "rigidity" as another implicit domain property of origami. When examining earlier literature of DNA origami, there are few studies testing the flexibility of 2 dimensional DNA, as often the origami was viewed using atomic force microscopy on flat mica. However, floating in 3 dimensional solution, origami potentially has a more dynamic nature (Failure VII of Appendix A). The researcher eliminated flexibility as a potential obstacle from the Slider by converting the *2D* structure to a *3D* one and by using feedback from CanDo, a dynamical tool for estimating flexibility in helix bundles shown in Castro et al. (2011) and Kim et al. (2012). The second iteration of the Slider, costing approximately $3000, achieved the fulfillment of a "rigidity" goal shown as LL#2 in Figure 2.4.

Between the second iteration and the third iteration, CanDo detected another obstacle to correct assembly of the intended structure, global twisting. Ke et al. (2009a) reported evidence on global twisting, as well as a method for correcting global twist. Dietz et al. (2009) manipulated the mechanism behind global twisting to bend origami on purpose. However, the researcher investigating the Slider had no knowledge of these two papers as global twisting only visibly affects *3D* origami (Failure VI of Appendix A). Again, *2D* origami is often deposited on flat mica, preventing analysis of global twisting. Similar to this experiment, Stein et al. (2011) found the obstacle global twisting by luck when examining a *3D* structure in CanDo.

In this case, the researcher independently found the relevant information in the literature. However, use of the failure catalog, which described this failure, could have helped find it as well. This find resulted in one less iteration of the project shown as LL#3 in Figure 2.4.

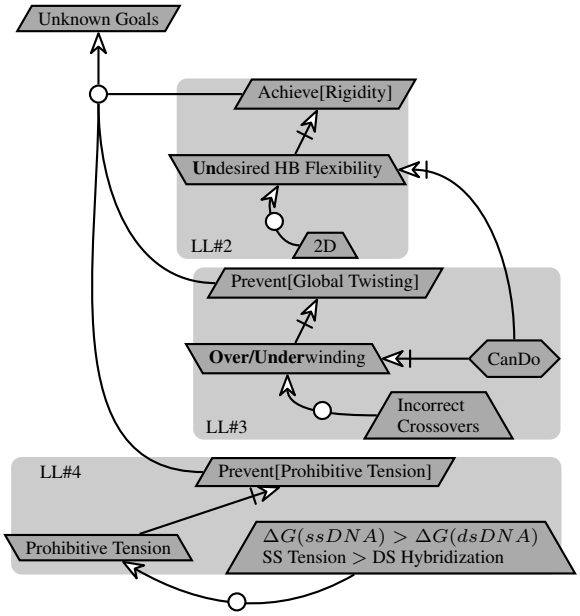


Figure 2.4 Updated Goal Model for the Slider

The third iteration still lacked the correct FRET output. After much experimentation the researcher for the Slider hypothesized that the tension of the single stranded domains in the scaffold prevented the correct formation of the origami, which also prevented the correct FRET signal during the analysis stage shown as LL#4 in Figure 2.4. Liedl et al. (2010) had conducted a study on single stranded DNA tension and origami, yet their paper did not explicitly label tension as a potential obstacle. Rather their study focuses on the construction of origami that can test the buckling limit of a specific structure containing rods composed of six helix bundles (Failure III of Apendix A). Similarly, the researcher could have used the failure catalog to either preemptively take the tension into account or diagnose the failure after detecting less FRET output than desired.

Many of the lessons learned from the Slider are common to other origami experiments. For example, in the experimental portion of the Laboratory for Algorithmic Molecular Programming, we now train rotation students about inter-origami stacking interactions, CanDo, and global twisting, as inter-origami separation, rigidity, and correct assembly are frequently goals

for our experiments. For the larger goal of greater efficiency in nanotechnology training, the catalog in Appendix A can help other researchers diagnose and prevent similar failures.

CHAPTER 3. A MODELING FRAMEWORK FOR DNA ORIGAMI WITH MULTIPLE CONFORMATIONS

DNA origami has emerged as a powerful tool for the creation of self-assembled, nanoscale devices. [1] Here I show how to develop DNA origami with multiple native states through a theory that describes DNA origami using mathematical models. Our models, the micro model and the macro model, define multi-state origami by creating equivalence classes of tertiary structures. Intuitively, if two different tertiary structures for one origami design are micro or macro equivalent, then both structures should form experimentally. To test this theory I designed a CaDNAno origami design that experimentally resulted in two differentiable tertiary structures. The resulting experiment successfully shows the formation of a purposefully designed DNA origami with multiple, kinetically trapped native states. I further tested the robustness of the internal stacking interactions of the structure through an "influencer" motif. The results, along with the design motifs, open up new strategies for creating DNA based devices with multi-state behavior.

3.1 Theory

The multi-state theory relies on converting potential tertiary structures for DNA origami into micro models and macro models in order to find "equivalent" origami structures. The micro model extends the DNA model presented by Winfree (1998) for defining the relationships between nucleotides. The macro model ignores the relationships between individual nucleotides, investigating instead the positions of crossovers which have similarities to the the polyhedra mesh design technique in Benson et al. (2015). Protein literature contain rigorous definitions for tertiary structures (McNaught and McNaught (1997)). Here, we define the tertiary structure

[1] The following contains work done in collaboration with Divita Mathur and Eric Henderson.

as an extension to the micromodel. It contains all of the information of the micro model, along with 3 dimensional coordinates for each nucleotide. A given DNA origami design has only one related macro model; however, the macro model will correspond to one *or more* micro models, and each micro model will correspond to one *or more* tertiary structures. Not all possible micro models result in energetically favorable tertiary structures, namely only micro models that represent the lowest approximate energy will appear experimentally.

To test the multi-state theory, I designed a 100nm×50nm DNA origami rectangle that uses strategically placed dumbbell hairpins for identification of state by atomic force microscopy (AFM). The origami structure had similarities to another multi-state structure found independently by Endo et al. (2014). The design for the origami rectangle places the dumbbell hairpins along its short edges. However, as predicted by the model herein, when implemented experimentally the design results in two distinct structures: the designed one and one with the same hairpins reoriented onto the long edges of the rectangle.

3.1.1 Micro Model

The micro model represents a DNA molecule as a graph of vertices and edges. Each vertex represents a nucleotide and edges represent relationships between nucleotides (Figure 3.1 *C*,*D*,*E*). Derived from the DNA model in Winfree (1998), a directed edge between two vertices represents a phosphate backbone in the 5' to 3' direction. An undirected (blue) edge represents a Watson-Crick bond. In the micro model, we add dashed edges to represent coaxial stacking between adjacent nucleotides (Yakovchuk et al. (2006)). We also label the vertices with colors to differentiate between scaffold (blue) and staple (other) nucleotides. In DNA origami, a helix bundle consists of a rod composed of double stranded DNA, where each nucleotide is connected via either a phosphate backbone or coaxial stacking. Coaxial stacking edges let us represent helix bundles in the micro model. Each micro model can represent one or more tertiary structures (Figure 3.1 *A*,*B*), as the nucleotides do not carry any information about their positions. *Intuitively, the micro (as in small) model looks at all the "small" relationships between individual nucleotides.*

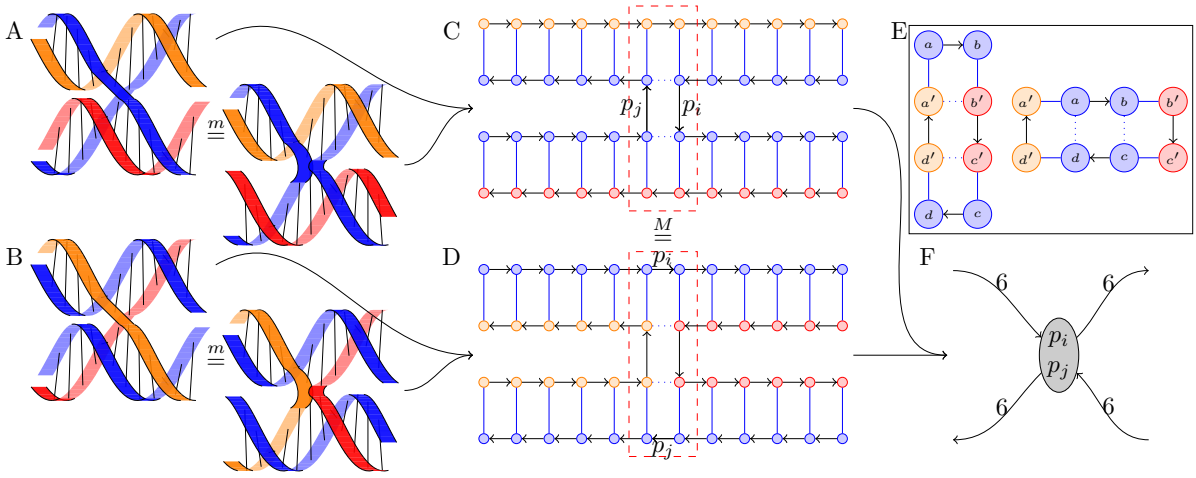


Figure 3.1 Tertiary, Micro, and Macro models for Origami. (A) Two tertiary structures with scaffold strands stacked (in blue). (B) Two tertiary structures with staple strands stacked (in orange and red). (C) Micro model corresponding to the tertiary structures in A. (D) Micro model corresponding to the tertiary structures in B. (E) Two possible coaxially stacked micro models for nucleotides a, b, c, d, a', b', c' and d'. (F) Macro model corresponding to micro models in C and D.

3.1.2 Macro Model

Each micro model corresponds to a macro model (Figure 3.1 F). The macro model focuses on crossovers (Figure 3.1 E) where a staple brings two disparate phosphate backbones together using adjacent Watson-Crick pairs along the staple (Appendix A). The phosphate backbones along the scaffold brought together by a staple uniquely label each crossover, encoded as a vertex. The directed edges between crossovers have weights equal to the number of nucleotides between each crossover. The direction of the edges corresponds to the $5'$ to $3'$ direction of the nucleotides. A solid directed edge represents a fully hybridized domain of DNA, while a dashed directed edge represents a partially hybridized domain. *Unlike the micro model, the macro (as in large) model looks at the big picture relationships between crossovers, but ignores small interactions between individual nucleotides.*

3.1.3 Micro/Macro Equivalence

Just as the micro models correspond to one or more tertiary structures, one macro model may correspond to multiple micro models (Figure 3.1 *C*,*D*) with alternative stacking conformations. As an example of a macro model, a single crossover taken out of the context of a larger DNA origami structure is an isomorph (graph-wise) of a Holliday junction shown in Chen et al. (1988). Each Holliday junction has two micro models, one for each possible stacking configuration. Additionally, when the crossover has four fold symmetry as defined by Chen et al. (1988) both stacking configurations form with equal likelihood (Figure 3.1 *E*, when $a' = b = c' = d$ or $a = b' = c = d'$). Then, depending on context, the micro models can correspond to multiple tertiary structures, such as antiparallel crossovers in antiparallel double crossover tiles or parallel crossovers in parallel double crossover tiles shown in Fu and Seeman (1993) or in unidirectional origami demonstrated by Han et al. (2013). Through the transformation of tertiary structures to micro models and micro models to macro models we can view tertiary structures in terms of equivalence classes (explained in Devlin (2003)). Two micro models are macro equivalent, $\overset{M}{=}$, if they correspond to the same macro model and have the same number of coaxial stacking edges. Two tertiary structures are micro equivalent, $\overset{m}{=}$, if they correspond to the same micro model and have similar free energies.

Given a DNA origami design and corresponding macro model, if there exists more than one micro model or more than one tertiary structure, then the predicted structures corresponding to each tertiary structure should appear in solution. However, not every macro model has multiple micro models, as nucleotides must have a specific orientation in order to coaxially stack. Over winding and under winding between crossovers breaks potential π stacks. Design motifs help isolate a small section of a larger DNA origami structure to elucidate the interaction between crossovers and over/under winding. A *motif* here describes any *small* (dependent on context), *repeating* substructure that when combined with other motifs forms a DNA nanostructure. Design motifs such as the *Odd Loop*, *Even Turn Loop* and *Seam Motif* force origami into one unique configuration (Figure 3.3-3.6). However, if an origami design lacks motifs that force one specific conformation, then (I hypothesized) it will have multiple corresponding micro

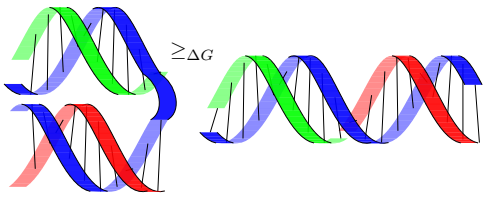


Figure 3.2 Coaxial Stacking

models and therefore, multiple tertiary structures. In other words, by necessity, non multi-state origami (with only one tertiary structure) must have a design that contains at least one or more multi-state preventing motifs, which will prevent multiple macro and micro models.

3.2 Design Considerations

3.2.1 Design Motifs

All of the following motifs use a blue strand to represent a segment of the scaffold in a nanostructure formed by the DNA origami method.

For simple DNA structures such as the duplex with one backbone nicked in Figure 3.2, there exists only one clear conformation with minimal energy. In this case, the coaxially stacked rod conformation has a lower minimal energy than the unstacked conformation due to the hydrophobicity of the nucleic acids. One can use thermodynamic parameters (Peyret (2000)) to calculate the difference in minimal energy (SantaLucia Jr and Hicks (2004)). For motifs belonging to a larger structure, we call the conformation with the minimal energy for the motif the *local* lowest energy conformation.

Figure 3.3 extends the example in Figure 3.2. Due to the helicity of DNA, the *Odd Loop Motif* only occurs when the loop connected to crossover $p_i p_j$ contains an odd number of turns. There are $\sim$ 10.4 nt per turn, thus the loop contains around $10.4 \cdot (2n + 1)$ nt. For such a loop, at least one pair of nucleotides cannot be stacked such that the DNA can bend to form a loop. Thus, any motif that has only one pair of nucleotides unstacked will have the local lowest energy. The *Odd Loop Motif* prevents multi-state structures by forcing the helix bundles extending from it (bound to the red strand) from falling into alternate conformations. Similar

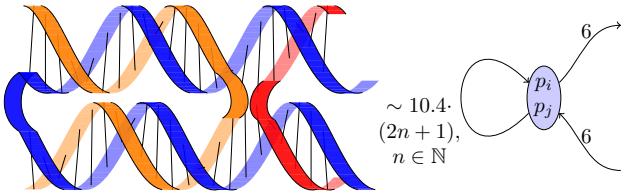


Figure 3.3 Odd Loop Motif

to Figure 3.2, the helix bundles will have a lower energy when they stack with the staple strand that composes the loop, making any other conformation energetically less favorable.

The *Even Loop Motif* in Figure 3.4 examines a loop that extends from an adjacent strand (attached to the green staple). Due to the helicity of DNA, at least two pairs of unstacked nucleotides are required to maintain an even loop structure. To create the unstacked pairs, the scaffold must bind with at least two domains of DNA (the red and orange strands). This motif prevents the crossover created between the scaffold, the red, and the green strands from having more than one energetically favorable stacking conformation. Thus this motif will result in only one valid micro-model; however, one can rotate the helix bundles attached to the red and green staples relative to each other to create more than one tertiary structure.

Specifically, the crossover will always be scaffold-stacked in the local lowest energy conformation. To show this, assume for contradiction that the red and green strands are stacked. The staple-stacked minimal energy structure would create the *Odd Loop Motif*; however, our loop has an even number of twists. Due to the helicity of DNA there does not exist an energetically favorable conformation for such a loop with an even number of twists, and we have reached our contradiction. The crossover will not be staple-stacked in the local lowest energy conformation. (It may however be possible that more than one motif attach to each other in such a way that none of them will stack in a locally lowest energy conformation.)

The *Seam Motif* in Figure 3.5 is a common motif found in an origami structure that has staples crossing a seam. It contains two parallel helix bundles held together by staples at energetically favorable (approximated by Rothemund (2006)) staple-stacked crossovers ($p_{i_0}p_{j_0}$ and $p_{i_2}p_{j_2}$), with a scaffold-stacked crossover in between ($p_{i_1}p_{j_1}$). Notice that on one side of the loop, the staple crosses over while on the other the scaffold crosses over. The *Seam Motif*

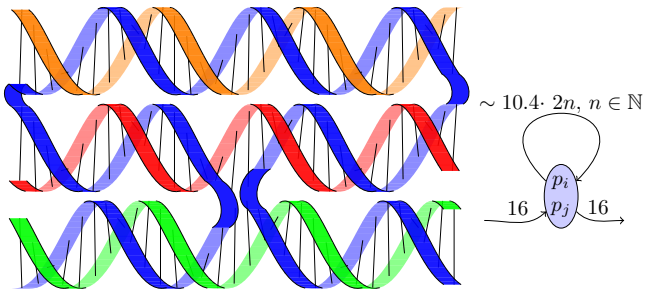


Figure 3.4 Even Loop Motif

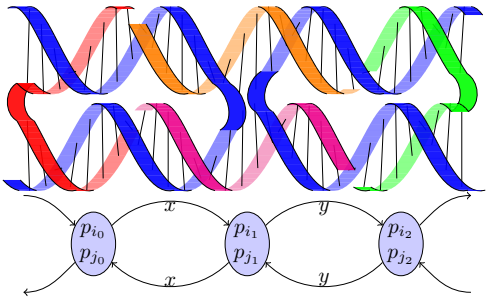


Figure 3.5 Staples Crossing a Seam, aka Seam Motif

essentially combines two *Odd Loop Motifs*. Thus, $2x = 10.4 \cdot (2n+1)$ and $2y = 10.4 \cdot (2m+1)$ for some $n, m \in \mathbb{N}$. The two odd loops force the crossover $p_{i_1}p_{j_1}$ into the scaffold-stacked conformation. Antiparallel, odd, double crossover tiles (Fu and Seeman (1993)) have a similar structure to the *Seam Motif* if we remove crossover $p_{i_2}p_{j_2}$ and ensure that there are two staples flanking $p_{i_0}p_{j_0}$.

If instead we force $2x$ and $2y$ to be an even number of turns, then the crossover $p_{i_1}p_{j_1}$ will have to be staple-stacked. However, only the coaxial stacks between staples in the helix bundles enforce this conformation. One could potentially squeeze the structure, forcing the crossovers to take a scaffold-stacked conformation to make a "loaded spring."

All of the motifs described so far permit only one conformation. While one can make multi-state origami by avoiding all of these mono-state motifs, it is also possible to create multi-state motifs out of mono-state motifs by inserting additional crossovers. The *Multi-State Odd Loop Motif* in Figure 3.6 is the *Odd Loop Motif* with the addition of the crossover $p_{i_1}p_{j_1}$. The nucleotides between the crossovers $p_{i_0}p_{j_0}$ and $p_{i_1}p_{j_1}$ $(y+x)$ must be an odd number of twists. By adding a nick along the backbone of the red staple strand (creating the green strand),

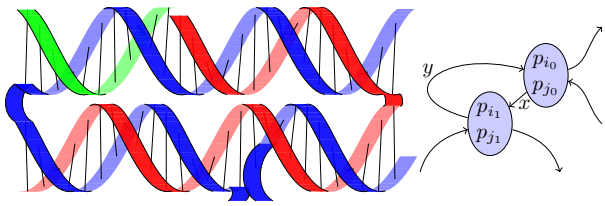


Figure 3.6 Multi-State Odd Loop Motif

the two crossovers $p_{i_0}p_{j_0}$ and $p_{i_1}p_{j_1}$ can be either scaffold or staple-stacked (in one of two conformations). When one is staple-stacked the other will be scaffold-stacked and vice versa.

A motif defines a small, repeating substructure that can combine with other motifs to form a DNA nanostructure. The motifs presented here do not represent a comprehensive list. However, they do reveal common design techniques used unintentionally that lock a structure into a desired shape. It is plausible that one might design a structure expecting one native conformation, only to find different structures by accident if they didn't employ any mono-state motifs. Conversely, the *Multi-State Odd Loop Motif* exemplifies how one can purposefully replace mono-state motifs to give a nanostructure multi-state behavior.

3.2.2 Design of Experiment

To test our models, I designed a DNA origami structure with multiple macro equivalent micro models (Figure S1-2), referred to as the rectangle design due to its shape. The macro model for the structure consists of a square lattice of a repeating multi-state motif: crossovers positioned approximately 1.5 helical turns, 16 nucleotides, apart (Figure 3.7 *A*). I purposefully avoided the *Odd Loop*, *Even Loop* and *Seam Motifs*, as any one of these motifs would remove the desired multi-state behavior. Each crossover has two possible stacking conformations: one where the scaffold coaxially stacks with itself (Figure 3.7 *B*, scaffold-stacked) and one where the staples coaxially stack (Figure 3.7 *C*, staple-stacked). The stacking conformation of one crossover forces adjacent crossovers to stack similarly (Figure S1-2). This results in two macro equivalent micro models each with a corresponding DNA origami design (Figure 3.7 *B*,*C*). Using CaDNAno (Douglas et al. (2009b)), both designs generated an identical set of staples (Appendix B Table S1), i.e. the staples for one design could also make the other design without

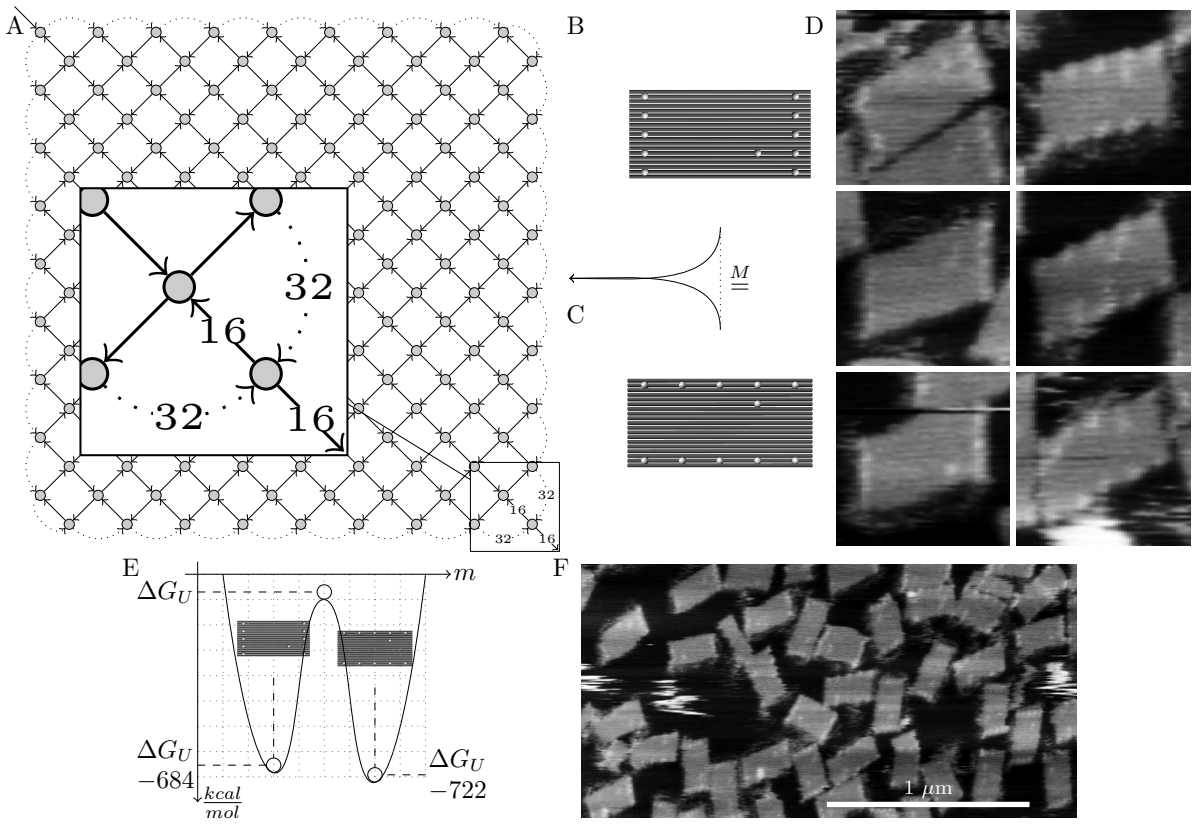


Figure 3.7 Multi-State Origami Experiment. (A) The macro model for the multi-state origami structure. Each gray circle is a crossover. (B) Cartoon for the scaffold-stacked conformation with dumbbell hairpins (gray circles). (C) Cartoon for the staple-stacked conformation with dumbbell hairpins. (D) AFM images of the scaffold-stacked (left) and the staple-stacked (right) conformations. (E) Relative free energies from an unstacked state ΔG_U (middle) to scaffold-stacked (left) and to staple-stacked (right). (F) AFM image of the multi-state structure experiment.

the need for extra strands or any other manipulations. I then labeled the DNA nanostructure with dumbbell hairpins using the same method as Rothemund (2006) and Woo and Rothemund (2011) along the edges (Appendix B Table S1) for AFM imaging. I also added an eleventh hairpin (Figure 3.7 B,C, and 3.9 A,D) to make the rectangle chiral mitigating Failure I of Appendix A.

The experiment's structure closely resembles studies of proteins with multiple native states, such as bacterial luciferase studied in Sinclair et al. (1994). Instead of creating an ensemble of conformations that the structure transitions freely between, our structure falls into one of

two states. Each state forces all of its crossovers into one of two stacking formations: scaffold-stacked (Figure 3.1 *C*, Figure 3.7 *B*) or staple-stacked (Figure 3.1 *D*, Figure 3.7 *C*). Since all of the crossovers must simultaneously stack one way or another, there exists a kinetic barrier exponentially proportionate to the sum of the coaxial stacking energies separating each state(Data S1-2). When computing this energy barrier I made the assumption that all staple domains bind successfully with the scaffold. Without assuming this, calculating the energy barrier would be **NP**-complete (Maňuch et al. (2011)) unless we only used two types of nucleotides (Mathieson and Condon (2015)). More experimental work, such as the one by Endo et al. (2014), has analyzed the energy barrier by parameterizing the annealing rate.

As for the multi-state behavior, Dunn et al. (2015) also created multiple states, but their method required two repeating m13mp18 sequences for their structure's scaffold, quadrupling the valid number of ways a staple can attach. Unlike their work, I leveraged weak coaxial stacking interactions instead of duplex formation to achieve multiple states, which removes the requirement for an altered scaffold. To demonstrate this, I used the same bacteriophage genome presented in Rothemund (2006).

3.3 Results

Upon annealing the staple and scaffold strands in solution, both the scaffold-stacked (Figure 3.7 *D*, left) and the staple-stacked (Figure 3.7 *D*, right) conformations formed, viewed by AFM (Figure 3.7 *F*). Calculating the thermodynamic free energies using nearest neighbor parameters as shown in Peyret (2000) and explained in SantaLucia Jr and Hicks (2004), I expected to see a bias towards the staple-stacked configuration (Figure 3.7 *E*, Data S2). As part of counting we made note of the relative orientation of each rectangle in order to identify potential interactions between the rectangle and the AFM probe as shown in Failure II of Appendix A. Upon counting the structures (Appendix B Figure S3-4, Table S3, Data S3) we found a yield for rectangles of 96.7% (standard error, $\sigma = 1.66\%, n = 3$). Of those with viewable hairpins ($92.7\%, \sigma = 3.74\%$), 32.4% of the structures stacked along the staples while 67.6% stacked along the scaffold ($\sigma = 1.66\%$). From these percentages, we conclude that rectangles of both conformations formed in solution.

Thus, the experiment verifies the existence of multi-state DNA origami nanostructures that fall into one of two states. It also provides evidence supporting the hypothesis that the tertiary structures of macro equivalent micro models will form experimentally in a solution. The experiment results showed a bias towards the scaffold-stacked conformation. We hypothesize that a strong bias towards the scaffold-stacked conformation emerged from global twisting created by the periodic crossovers as shown in Failure II of Appendix A (Appendix B Figure S5). Removing global twisting while ensuring multiple states is not trivial and remains as future work.

We also investigated the orientations of rectangles to see if the AFM created a bias by interacting unfavorably with one structure and not the other. Counting rotations of the staple vs. scaffold-stacked conformations showed a similar distribution for each (Figure 3.9 *C*) with no evidence of interaction with the AFM other than a bias to be oriented 30° above the horizon. Counting whether dumbbell hairpins were face up or face down did show a clear bias for the staple-stacked conformation to lay face up on the mica ($83.5\%, \sigma = 2.8\%, n = 8$). With a sample as low as 14 rectangles, using a binomial test determines the likelihood for a staple-

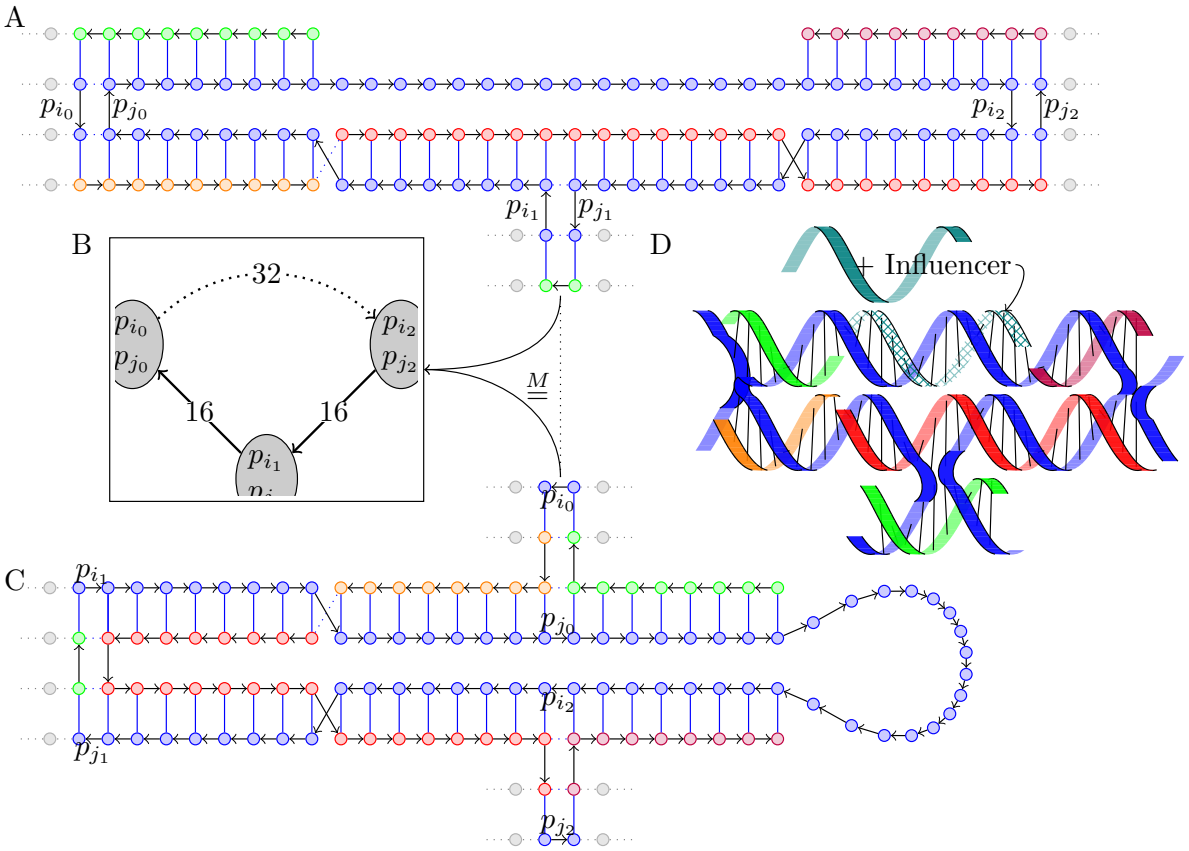


Figure 3.8 Influencer Motif for Final Configurations. (A) Micro model of scaffold-stacked conformation. (B) Macro model corresponding to micro models A and C. (C) Micro model of staple-stacked conformation. (D) Tertiary structure for scaffold-stacked conformation with an influencer strand. The influencer strand causes a lower well of potential energy for the scaffold-stacked conformation than for the potential energy for the staple-stacked.

stacked rectangle to land fairly on either side to be only 0.65%. This correlation indicates global twist likely influences which side of the staple-stacked conformation binds with the mica (Data S2, Figure S5).

The long edges of both the staple-stacked and scaffold-stacked states have 18 *Influencer Motifs* (Figure 3.8 *B*). Each motif has two macro equivalent micro models: scaffold-stacked (Figure 3.8 *A*) and staple-stacked (Figure 3.8 *C*). Without a sequence-specific influencer strand (Figure 3.8 *D*, Table S2), either conformation can form experimentally. Theoretically, the addition of an influencer strand (Figure 3.8 *D*) should encourage the formation of one micro model (the scaffold-stacked, Figure 3.8 *A*) over another (the staple-stacked, Figure 3.8 *C*). However, initial results show that the ratio of both of the structure's states is fairly resistant to the addition of a few influencer strands (< 6, Fig 3.9 *B*). This suggests that multi-state structures can form even with a relatively small number of mono-state motifs. A few mono-state motifs do not seem to have enough influence to stop the multi-state behavior of our structure.

3.4 Discussion

In this work, I have reported the simultaneous formation of two different nanostructures from one CAD DNA origami design. Our experiment, which also tests the folding thermodynamics of an influencer motif, shows an example of a multi-state DNA origami using the m13mp18 scaffold. Our macro and micro models, along with the macro and micro equivalence theories, explicate the existence of multi-state nanostructures and provide a conceptual framework for how to design future multi-state structures and motifs. Future work expanding these models should aid CAD software in the construction of nanodevices incorporating motifs, opening up a new way to conceptualize the design of DNA nanostructures. The experimentally verified, multi-state motif expands the already existing repertoire of DNA origami design strategies. Unlike the work by Endo et al. (2014), the rectangle structure had the multi-state property by design according to our theory. Both experiments together give insight into more general multistate designs.

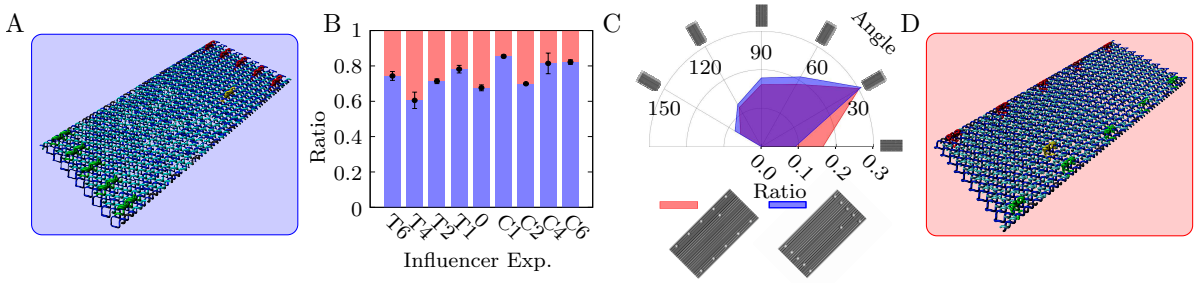


Figure 3.9 Results of Multi-State Origami Experiment. (A) Designed tertiary structure for scaffold-stacked conformation. (B) Normalized ratios of observed staple vs scaffold-stacked conformations for influencer experiments (C6 = 6 scaffold-stacked conformation influencers, ... 0 = no influencers, T1 = 1 staple-stacked conformation influencer, etc). (C) Normalized ratios of observed staple and scaffold-stacked conformations per rotation from the horizon. Rectangles found to be -30° were counted with rectangles found to be 150° ($180° \equiv 0°, 210° \equiv 30°$, etc). ($D$) Designed tertiary structure for staple-stacked conformation.

Integrated DNA Technologies (IDT, Coralville, IA) supplied all oligonucleotide staple strands in RNase free water at 100μM concentration in individual wells. Bayou Biolabs (Matairie, LA) supplied m13mp18 single stranded scaffold strand DNA at a concentration of $1\frac{\mu g}{\mu L}$ in Tris -Acetate EDTA buffer. Fisher Scientific supplied Tris-Acetate EDTA, Magnesium Acetate Tetrahydrate, and water.

The assembly protocol for the rectangle structure followed the protocol originally laid out by Rothemund (2006) designed using CaDNAno (Douglas et al. (2009b)). The staple strands each had a final concentration of 50 nM, mixed with the m13mp18 scaffold strand at a final concentration of 10nM in 1x reaction buffer (comprised of 40 mM Tris-Acetate, 1 mM EDTA (pH 8.3) and 18 mM Mg^{2+}) and brought to a final volume of 50 μL. The influencer strand experiments followed the same protocl, with each influencer mixed in at a final concentration of 50 nM. Images were generated using Tapping Mode^{TM} (with AC40 AFM probes from Bruker) on a Digital Instruments MultiMode AFM. 5 μL of structure mix was deposited on freshly cleaved mica and directly imaged. Images were corrected to account for stretching artifacts caused by AFM. Three researchers independently counted structures from the AFM output to generate all statistics shown.

CHAPTER 4. SUMMARY AND DISCUSSION

The complexity of creating optimal DNA origami designs rivals the most difficult problems in Computer Science (it is *NP*-complete, Ellis-Monaghan et al. (2014)). Current techniques can create arbitrary shapes, such as a dolphin (Andersen et al. (2008)), a tetrahedron (Ke et al. (2009b)) and an opening and closing box (Andersen et al. (2009)). However, the current literature lacks models for multi-state origami. Such models may give insights into the kinetics of how origami assembles (Endo et al. (2014) and Dunn et al. (2015)). Creating functional origami (Kuzuya and Ohya (2014)) requires both the correct assembly of DNA origami and the implementation of dynamic behavior. For both the assembly and implementation tasks, there exist several known failures and errors in design. Work on distilling the literature and personal experiences of lab scientists can help future researchers define specifications for their DNA nanostructures so that their devices work correctly and/or with improved yield.

Chapter 1 introduces DNA origami and the relevant related works. Chapter 2 then shows the creation of a failure catalog available in text form in Appendix A. Chapter 2 also shows the application of the catalog to a real world case study named the Slider. Each of the lessons learned from its iterative development update its goal model, creating new requirements after each iteration of the experiment. Chapter 3 approaches requirements gathering from a modeling perspective. Chapter 3 details the creation of a novel model with enough explanatory power to create multi-state origami. It also describes the use of the failure catalog to preempt certain failures in a real world experiment to verify the model. Appendix B shows how to recreate the experiment, along with more specifics about the creation of the model, the analysis of the data, and the preference of the scaffold-stacked vs. the staple-stacked state.

Chapter 3 introduces the micro and macro model together with motifs that allow or prevent multi-state behavior in section 3.2. Chapter 3 then shows a DNA origami design that avoids

all motifs preventing multi-state behavior, providing evidence for the correctness of the micro and macro models. Future work can extend the use of motifs as a way to abstract the routing of the scaffold and placement of staple strands. DNA origami may then be able to use motifs from a pattern catalog in a way similar to how one uses design patterns in software engineering. The macro model also gives a novel way to analyze the formation and possible states of DNA origami. Future work with the models can use the minimal cycle basis (Amaldi et al. (2010)) of the macro model to see the underwinding and overwinding of a cycle depending on placement and configuration of crossovers. Other interesting open problems include analyzing the conformation of individual crossovers using a coarse model such as cellular automata, and the creation of origami with more complex crossovers (with 4 incoming edges and 4 outgoing edges as opposed to 2 incoming edges and 2 outgoing edges).

By attacking the design problems intrinsic to DNA origami using prior experience, through a failure catalog, and using theory, through macro and micro models, this work provides new insight into how to avoid problems when designing future experiments.

APPENDIX A. DNA ORIGAMI FAILURE CATALOG

The following catalog went through review and revision conducted with Divita Mathur and Robyn Lutz.

A.1 Definitions

D1.Anneal: When one heats and then reduces the temperature of a solution over a period of time to aid in the synthesis of a product, we call it annealing. Annealing can increase yields; however, new research and better understanding about the formation of DNA origami may eliminate this step (Sobczak et al. (2012)). Some researchers also refer to annealing as the process of bringing two strands together, similar to hybridizing.

D2.Antiparallel: When two helix bundles are parallel to each other, each helix bundle will have two single stranded DNA: one going from the 3' to the 5' end and one going from the 5' to the 3' end (in reverse). Thus a pair of the single stranded DNA between the two helix bundles in parallel goes in opposite directions (3' to 5' and 5' to 3'), and we call this antiparallel. Specifically, in DNA origami, the rastering of the scaffold orients it such that the scaffold is always antiparallel with itself in adjacent helix bundles connected by crossovers. See Chen et al. for an example of anti-parallel helices (Chen et al. (1988)).

D3.Atomic Force Microscopy: An atomic force microscope uses a probe (also known as a tip or cantilever) to scan the topography of small molecules with 5nm resolution. The

result of this operation is a height map corresponding to the height at which the needle interacted with the molecule of interest.

D4.Base Pair: When two single stranded DNA domains hybridize (come together) their individual nucleotides (bases) connect according to Watson and Crick's theory for pairs (A with T and C with G). One such pair is called a base pair. Base pairs are also a measurement of length for double stranded DNA.

D5.Blunt Ends: When a segment of double stranded DNA ends, exposing one face of the hybridized base pair unstacked, we call this a blunt end.

D6.Blunt End Stacking: Between base pairs there exists a force that pulls the bases closer together. The hydrophobicity of the base pairs (their desire to avoid water molecules) causes them to "stack." Thus when blunt ends come together they will stack next to each other. With DNA origami this effect noticeably causes structures to align along their blunt ends, creating long chains of DNA origami.

D7.CaDNAno: Software that allows for the rapid prototyping of DNA origami (Douglas et al. (2009b)).

D8.CanDo: Software that takes a file from CaDNAno and approximates the resulting origami's rigidity, global twisting, and conformation. See Kim et al. (2012) for an explanation of how the software works.

D9.Cation: A positively charged atom or molecule. See the definition for salt.

D10.Chirality: A molecule is chiral if it is unique from its mirror image. Conversely, an achiral molecule is not unique, e.g. in origami, an achiral structure looks the same when you flip it or orient it a different way.

D11.Coaxial Stacking: When two nonadjacent DNA domains hybridize with two adjacent domains along a single stranded DNA (ssDNA). In DNA origami, adjacent staples have coaxial stacking. See SantaLucia Jr and Hicks (2004) for some examples.

D12.Complementary: A nucleotide, A, is complementary to another nucleotide, B, if the unordered pair (A,B) is equivalent to either (A,T) or (C, G).

D13.Codomain: A codomain, A, is a domain whose sequence of nucleotides are reversed and complementary to another domain, B.

D14.Crossover: When the backbone of one single stranded DNA jumps from one helix bundle to another, we call it a crossover (for use of the word "jumps" see Woo and Rothemund (2011)). Crossovers can be made with either the scaffold or the staple strands.

D15.Domain: A domain of DNA is a specific, continuous sequence of single stranded DNA.

D16.DNA Origami: A method for creating nano structures by folding a long length of DNA called a scaffold. See Rothemund (2006).

D17.Electron Microscopy: Using an electron beam, a small object is imaged either through the transmission of the electron beam or through the conversion of the beam into other forms of energy such as heat. In transmission electron microscopy, the molecules to be observed often need to be dehydrated. Single particle cryo-electron microscopy does not require dehydration, but it does require the solution to be frozen.

D18.Fluorescence Microscopy: Fluorescence microscopy is a way of analyzing DNA strands with light emitting fluorophores in such a way that the appearance or disappearance of visible light informs one on the state of an experiment. Techniques such as Förster resonance energy transfer (FRET), total internal reflection fluorescence (TIRF) and fluorescence quenching use fluorophores, and are all techniquess used in fluorescence microscopy.

D19.Gel Electrophoresis: A process where macromolecules (such as DNA) are put into a gel with an electric current. The result is the filtration of DNA molecules (including hybridized DNA) by their size in different parts of the gel as they migrate in presence of a voltage difference. One can use this to separate complete DNA origami from materials, the ratio of which represent the yield.

D20.Helix Bundle: A helix bundle (hb) is a consecutive section/collumn of antiparallel double stranded DNA.

D21.Inter-helical Gap: Two adjacent helix bundles connected by crossovers will repel each other such that there exists a gap between the two helices. We call this the inter-helical gap.

D22.Nucleotide: A nucleotide (nt) is one base, sugar and phosphate of a single stranded DNA. It can have one of four possible bases, A, T, C or G, which bind according to Watson Crick interactions.

D23.Overwound: When nucleotides are deleted between ideal crossovers (every ~ 10.4 nt), each nucleotide must extend its twist per nucleotide. This results in less than 10.4 nt per twist causing each nucleotide to be overwound. Mathematically, examine a full twist of 10.4 nt for a domain A versus a full twist of 9 nt for a domain B. Domain A's nucleotides have a $\frac{1}{10.4}$ twist whereas B's nucleotides have a $\frac{1}{9}$ twist. Thus B is overwound by $\frac{1}{9} - \frac{1}{10.4} \approx 0.15$.

D24.Polarity: Just as atoms have positive or negative charges, molecules will have polarity, a vector from negative to positive of a change in charge over a length of the molecule. Polarity largely affects hydrophobicity (retraction from water), as hydrophobic molecules are non-polar and hydrophilic molecules are polar.

D25.Salt: For creating DNA origami in solution, the solution must be composed of water and a concentration of counterions. Most commonly DNA nanotechnology will use NaCl, MgCl2, MgOAc or NICl2 to provide cations Na+ or Mg2+, but there are many other

options as well. The salt will bind with the backbone of the DNA, reducing its polarity. With too much polarity (too little salt), DNA is unable to bind.

D26.Scaffold: The scaffold for DNA origami consists of a long single stranded DNA usually > 7000 nt in length, which is folded to create a desired structure. Often, DNA nanotechnology uses the genome of the m13 bacteriophage; however, other long single stranded scaffolds could also work.

D27.Seam: In DNA origami, a seam is a contour of the origami for which the scaffold strand does not cross. Seams can be closed by crossing the contour with staple strands.

D28.Seed: A seed domain is a domain of a staple that binds with a codomain of the scaffold that is 14 nt in length, as presented by Ke et al Ke et al. (2012). However, the length, 14nt, should not be considered as a necessary constraint for a seed domain, as any domain that binds with high specificity before other staple domains should theoretically work the same way as a seed.

D29.Single-Layer 3D Origami: 3D origami that is comprised of flat polygons (made from anti parallel helix bundles), constructing the surface area of an object, e.g. a hollow cube. See Linko and Dietz (2013) for more examples.

D30.Solution: Solution is generaly a specific ratio of salt mixed with water.

D31.Staple: A staple is a single stranded DNA that binds two disparate parts of a scaffold strand. The two disparate domains are held in place by their domains and the backbone of the staple.

D32.Strain: The strain on a segment of DNA can be thought of as the potential energy created by deforming the DNA idealized as a spring from its lowest energy state. Think of strain on DNA as the squeezing, stretching, twisting, and bending of a double helix.

D33.Strand Displacement: When part of one dna strand binds to another single stranded domain, allowing it to "invade" adjacent hybridized domains. See Qian and Winfree (2011) for more information.

D34.Transmission Electron Microscopy: The dehydrated version of electron microscopy. See the definition for electron microscopy.

D35.Underwound: When nucleotides are inserted between ideal crossovers (every 10.4 nt), each nucleotide must shorten its twist per nucleotide. This results in more than 10.4 nt per twist causing each nucleotide to be underwound. See the definition for overwound for calculating under or over winding.

D36.Walker: Molecular motors, such as kinesin and dynein, transport molecular payloads in alive cells. Similar attempts have been made with DNA nanotechnology. Most move a strand of DNA using strand displacement such that the motor moves with a bias in one direction along a track (also known as a substrate since the walker interacts with it).

D37.Yield: The yield is the fraction of correctly formed DNA origami over the total amount possible. This can be measured using gel electrophoresis. Researchers will some times rely on visual analysis via AFM to determine yield; however, this method introduces observer bias and sampling error.

A.2 Failures

I.2D Chirality

Assumed Properties: Symmetry, Chirality, Reference Marker

Affected System Goals: Analysis, Observation Bias

Failure Type: Compound

Description: Though 3D origami can also have issues with chirality, most notably asymmetrical 2D origami used for walkers may convince researchers of false confirmations depending on the the orientation of the origami (face up vs face down).

Origin: An origami dolphin shows proof of concept that 2D origami can land with approximately equal probability on either of its two faces (Andersen et al. (2008)). Lund et al. (2010) used chiral origami for a DNA spider walker that followed a track from start to stop. They discovered that when the origami was face down on the mica, people were more likely to identify a spider in the stop position, even when there was no walker. Another walker presented in Wickham et al. (2012) that had to navigate a symmetrical set of tracks along a 2D origami was given a "reference marker" in the lower right hand corner to give the structure visible chirality. Jungmann et al. (2014) found in an experiment resulting in digital numbers displayed by FRET based DNA paint strands that the digit "5" was showing sometimes when "2" was intended. It is assumed that the "5" was actually a "2" flipped face down.

Cause: When depositing 2D origami on a surface to be observed it can with approximately equal probability land on either side. This can lead to false observations or unexpected behavior.

Detection: origami that is absolutely not chiral will land on either side with equal probability, and it will not be possible to tell which side it is on. Also, if results of a chiral origami experiment seem to correlate with one side and not the other, then the orientation of the origami with respect to the observer likely matters.

Solution/Mitigation: For non-chiral origami, one should implement chirality in design. For example, reference markers let the observer (researcher) distinguish between which paths the walker was on (Lund et al. (2010)). Next, to distinguish between whether the topside (side facing the observer) or the bottom side are giving false observations, one can implement a double blind experiment (Wickham et al. (2012)) where the real outcome is known. Thus

one can compare the expected outcomes from either orientation of the origami, comparing them both with the real outcome to get an estimate of the bias associated with each.

II.AFM Distortion of Origami

Assumed Properties: AFM (tip, cantilever), Single-Layer 3D Origami (Hollow 3D Origami), Higher Resolution (higher force on tip)

Affected System Goals: Analysis, Feedback

Failure Type: Component

Description: Atomic Force Microscopy returns either an altered structure, or it damages the structure.

Origin: Found when examining a DNA origami dolphin. The tail of the dolphin was distorted when applying extra force on the AFM tip: "this tail distortion can clearly be assigned as an influence of the force exerted by the AFM tip" (Andersen et al. (2008)). Furthermore, the direction of scanning affected the dolphin's tail. In several 3D origami shapes with hollow cores (Ke et al. (2009b), Kuzuya and Komiyama (2009) and Andersen et al. (2009)), the AFM tip caused the shapes to either be shorter than expected (Kuzuya and Komiyama (2009)) or caused the shapes to completely collapse (Ke et al. (2009b) and Andersen et al. (2009)). Rothemund (2006) found that 26-helix squares were being stretched by the tip of the AFM creating hourglass shapes. Furthermore, in labeled staple experiments (where the staple formed a hairpin), the AFM may have also damaged the labeled staples (Rothemund (2006)).

Cause: Atomic Force Microscopy works by scanning a 3 dimensional surface with a needle in order to create a height-map corresponding to the surface. (Note, this does not mean it

only works for 3D origami. 2D origami also has a height). The force applied by the needle touching an origami can distort or destroy the intended structure of the origami.

Detection: Increasing the resolution for an AFM, which also increases the force, may create noticeable distortions (Andersen et al. (2008)). Differences in design and AFM output, especially differences in height or in one direction, imply an unwanted interaction with the tip. Sequential imaging can show deformations that occur over time, which may suggest unwanted interactions with the needle (Rothemund (2006)).

Solution/Mitigation: Use a different imaging technique such as Electron Microscopy or DNA paint (Jungmann et al. (2014)). The technique chosen will depend greatly on the design and cost constraints of the experiment. If an AFM is the only choice, then lowering the force or resolution should help.

III.Compression Limit for Helix Bundles

Assumed Properties: Single Stranded Domain, Tension, Base Pair (Hybridization), Compression

Affected System Goals: Assembly, Task

Failure Type: Compound, Component

Description: Origami does not form due to buckling under compression forces. A helix bundle will fold/remain folded with a compression force less than the critical Euler Force F_c (also known as the buckling force).

$$F_c = pi^2 * P * k_b * T/L^2$$

T = Temperature

k_b = Boltzmann constant

$pi = 3.14...$

P = Persistence length of bundle

L = Length at which helix buckles

Origin: Experimentally verified in tensegrity structures that formed kites with struts composed of 12-helix-bundles (12-hb) studied in Liedl et al. (2010).

Cause: In the original experiment the compression was caused by single stranded DNA (ssDNA). The ssDNA was tightened until the struts (in this case 12-hb) failed (they no longer formed the intended structure as shown by lower yields when examined by gel electrophoresis).

Detection: Detection is done primarily through gel electrophoresis, though the effect of buckling may be observable with electron microscopy.

Solution/Mitigation: Calculate the compression force exerted on the helix bundle in question (or all helix bundles) and reduce the cause of the force or adjust the length of the helix bundle (Liedl et al. (2010)). Information on the forces can also be found in an older study by Smith et al. (1996). Also, note that gel purification was found to lower persistence length which changes the calculation of the force (Liedl et al. (2010)).

IV.Denatured Base Pairs

Assumed Properties: Staple Domain (Hybridization), Base Pair (Hybridization)

Affected System Goals: Analysis

Failure Type: Cumulative

Description: For base pair by base pair (bp) conformations most imaging techniques are insufficient. While they can give an image of the general shape of the origami, one cannot tell which segments of the origami are denatured (unattached), if any.

Origin: A solution to this problem was addressed by Wagenbauer et al. (2014), whose technique can identify regions where 3 or more nucleotides are unbound.

Cause: Current techniques for showing whether or not intended domains of DNA strands have hybridized are based on imaging techniques such as electron microscopy (EM) or atomic force microscopy (AFM). Such techniques do not have bp resolution.

Detection: Denatured base pairs are hard to detect. Other than the method described in the solution, gel electrophoresis alone can show large amounts of denatured bps through bands corresponding to structures other than the origami. If a structure has near one hundred percent yield, this measure may not help.

Solution/Mitigation: To show how many base pairs are hybridized correctly, Wagenbauer et al. designed a circular de-Brujin sequence containing all domains of length 3. (In other words all permutations of A, C, T, G of length 3 were contained as a substring of the larger de Brujin sequence.) The de-Brujin sequence was then split from 64 nt domains into two smaller 47 nt and 27 nt domains to avoid secondary structures such as hairpins. These strands are the 'defect labels' (Wagenbauer et al. (2014)). In one's own experiments, Wagenbauer's method can be implemented in the following steps:

1. Attach a fluorophore to the defect labels.
2. Add the defect labels to the solution containing the origami in question.
3. Use gel electrophoresis to separate the denatured defect labels from the origami.
4. Compare the intensities of the fluorescence in each band to see how many of the defect labels have attached to the origami structure. See Wagenbauer et al. (2014) for a comparison of results.

Wagenbauer et al. used this information to fine tune the annealing ramp (cooling step) for 42 helix bundle (hb) 3D origami structure.

V.Electrostatic Repulsion

Assumed Properties: Crossovers, Underwinding, Overwinding, Inter-helix Gap, Crossover Spacing

Affected System Goals: Assembly

Failure Type: Component, Cumulative

Description: Having more nucleotides (nt) between crossovers results in a greater gap between helix bundles (hb) due to the repulsion. Underwinding can relieve strain created from the repulsion, while overwinding can decrease the distance between two strands, increasing the strain. The effect is similar to when one tries to push the negative ends of two magnets together.

Origin: Rothemund (2006) reported on this effect in his original paper on DNA origami. Specifically, he reported an "inter-helix gap of 1 nm for 1.5-turn [16 bp] spacing and 1.5 nm for 2.5-turn [26 bp] spacing." In multilayer origami with 7bp spacing the inter-helix gap shrinks to 0.5 nm (Ke et al. (2009a)). Finally Ke et al. (2012) found that underwinding staples to give 3D structures relief from electrostatic repulsion resulted in higher yields.

Cause: Electrostatic Repulsion is due to the polarity of the sugar backbone of DNA. The negatively charged backbone's hydrophilicity along with the nonpolar bases drives DNA to form; however, for two anti-parallel dsDNA strands the negative dipole moment of the sugars will also cause the strands to repel from each other.

Detection: For 2 dimensional origami the gap can be found by measuring the y-resolution of the origami (using AFM or EM). A similar technique can be used for 3D origami. For origami that has low yields, electrostatic repulsion could be part of the problem.

Solution/Mitigation: For origami that needs a specific inter-helical gap, reducing or increasing the space between crossovers will reduce or increase the gap. For 3D origami, targeted insertions which increase the base pairs between crossovers underwind the structure relieving tension created by electrostatic repulsion. This also can be interpreted as increasing the inter-helix gap. Furthermore, an intercalator such as ethidium, combined with underwinding can increase yields. (One should note that this result is design dependent, and the underwinding should be done such that the interaction with the intercalator results in an "equilibrium twist" of 360 degrees as shown in Ke et al. (2012).)

VI.Global Twisting

Assumed Properties: Crossovers, 3D origami, CanDo, CaDNAno, Overwinding, Underwinding, Square Lattice, DNA helix

Affected System Goals: Assembly

Failure Type: Component

Description: Origami has too much or unwanted torsion that affects structural functionality (Mathew-Fenn et al. (2008)).

Origin: It is found in experiments on 3D origami comprised of 2 and 3 layers, but not found in origami of 6 or 8 layers (Ke et al. (2009a)).

Cause: Local underwinding of helices results in global relaxation (Ke et al. (2009a) and Dietz et al. (2009)). In short, a global right-hand twist results from the strain of underwinding.

Note that the CaDNAno software (Douglas et al. (2009b)) designs square-lattice 3D origami with global twisting by default.

Detection: One can design the origami to form multimers such that the twist is elongated over multiple monomers. This makes the total twist easier to measure as a function of length (Ke et al. (2009a)). Computational methods, such as CanDo, for detection also exist (Kim et al. (2012)).

Solution/Mitigation: Change the initially imposed double helical twist density such that staples are overwound instead of underwound. One way to accomplish this is deletion of base pairs spaced evenly along the length of the scaffold (192-bp in the original paper). The experiment was carried out with 10.5 bp/turn (removing 3 bp to get (192 - n)/(24 * 0.75) = 10.5, n = 3), 10.44 bp/turn (removing 4, n = 4), and 10.39 bp/turn (removing 5, n = 5). No twist was observed for the 10.44 or 10.39 bp/turn designs (Ke et al. (2009a)). Reasons for why the staples had to be overwound are speculative, but the authors offer the work by Mathew-Fenn et al. (2008) as a possible explanation. Stein et al. (2011) also has a nice explanation of how to fix global twisting using CanDo.

VII.Helix Bundle Flexibility

Assumed Properties: CanDo, Coaxial stacking, Crossovers, Flexibility, 2D Origami

Affected System Goals: Task, Assembly

Failure Type: Component

Description: In solution a DNA helix bundle (hb) has a certain amount of flexibility that can either prohibit or enable the operation of a desired task.

Origin: Nanotechnologists involved in producing origami have been aware of, and have purposefully tried to utilize the flexibility of origami. A good example comes from Andersen et al. (2008) who tried to make a dolphin tail that wiggled.

Cause: Flexibility is an intrinsic property of DNA. Single stranded DNA has several degrees of freedom in its backbone. Double stranded DNA (dsDNA) reduces, but does not eliminate the aforementioned freedom through Watson-Crick base pairs (bp). In practice, double stranded DNA has a persistence length of 50nm, after which dsDNA acts like a flexible rod as opposed to a rigid beam.

Detection: Methods for detection in solution require examining the differences from different instances of the same origami design in solution using a method that can capture the physical formation of many DNA origami such as TEM. Also, the software utility CanDo (Castro et al. (2011) and Kim et al. (2012)) can predict a lower bound for the flexibility of a helix bundle caused by perturbations from the heat of the solution. It creates a heat map which corresponds to the movement (in nm) of a helix bundle (hb) in solution.

Solution/Mitigation: To increase rigidity, one can either change the placement and quantity of crossovers (increasing for 3D origami), or if the current DNA origami is 2D, then one can make their design 3D by adding layers.

VIII.Inter-Origami Base Stacking

Assumed Properties: Stacking, Exposed Blunt Ends

Affected System Goals: Assembly, Task, Analysis, Origami Separation, Individual Origami, Coagulation

Failure Type: Interaction

Description: When the ends of staples, also known as blunt ends, on two origami structures meet, they have a tendency to "stick" to each other. This can cause unexpectedly large chains of 2D origami or can be manipulated purposefully. See Douglas et al. (2009a) for an example of stacking with 3D origami.

Origin: Found by accident in the original paper on origami by Rothemund (2006), but also investigated and manipulated by Woo and Rothemund (2011). A good example of how to use base stacking to make a large structure is the stack cross (Douglas et al. (2009a)).

Cause: Each base of DNA, though non-polar, does have an electron cloud (π-cloud) that encourages bases to stack on top of one another (π-bond). This occurs even when nucleotides are not attached by a sugar-phosphate backbone.

Detection: Can be detected by atomic force microscopy (AFM) or potentially by fluorescence resonance energy transfer (FRET) if the sides of the origami are labeled. The edges of the origami are where the blunt ends will adhere to the blunt ends on other origami causing blunt end stacking.

Solution/Mitigation: To get rid of the effects of inter-origami base stacking, single stranded segments of the scaffold along the edges can prevent the effect. There are two main ways to accomplish this:

1. Add staples which cause T loops (single stranded domains of only T nucleotides connected to an edge of the scaffold, such that it forms a loop). A's cannot form loops due to their single stranded persistence length. G's cannot form loops due to the formation of quadruplexes.

2. Leave single stranded sections of the scaffold unbound on the edges, preventing the formation of blunt edges.

Woo and Rothemund (2011) encouraged blunt ends to form by:

1. Making all blunt end pairs "GC" instead of "AT", since "GC" blunt ends form strong pi bonds.

2. Correcting global twist. ("Chains of twisted origami [broke] with a characteristic offset" Woo and Rothemund (2011))

3. Preserve the B-form twist of the DNA, such that the major/minor grooves are aligned using "relaxed edges". Relaxed edges have the blunt ends of the staple on the edge of the origami lattice vs. the middle or inside of the staple on the edge.

IX.Remainder Strand Displacement

Assumed Properties: DNA Domain, Displacement, Multiple Unique Origami, Remainder DNA

Affected System Goals: Assembly, Yield

Failure Type: Interaction, Cumulative

Description: When multiple DNA origami with different underlying designs are mixed in the same solution, unbound remainder strands from one solution can invade and displace strands in a differently designed origami originally mixed in a separate solution. When a DNA origami design uses less than all of the nucleotides in a scaffold, the rest of the scaffold is sometimes filled with remainder strands to prevent unpredictable binding with the remainder single stranded domain.

Origin: In the supporting info. of Woo and Rothemund (2011) which involves multiple origami with different designs, they found that staples left floating in solution from one origami (remainder staples) would invade and unfold origami with a different design by attaching to

single stranded loops. Specifically, the remainder staples caused "large structural disruptions."

Cause: The different remainder strands from the different designs correspond to different domains along a scaffold. A domain that might be bound in one origami design will be unbound in another. Thus freely floating remainder strands will bind with these open domains when different origami designs are mixed. As Woo and Rothemund (2011) state, this process can continuously open up domains, and the newly formed double stranded DNA may be a more energetically favorable continuous duplex due to coaxial stacking (see their supporting info.). They did not notice a similar error with just staple strands.

Detection: Potential unwanted interactions can be pre screened by examining the staple sequences. In solution, the structural disruptions depending on severity should show up using an AFM.

Solution/Mitigation: Woo and Rothemund (2011) solved this problem by aligning the remainder loops such that the remainder domains and strands were identical. They note that it is also possible to not add the remainder strands altogether. One could also purify the different DNA origami with a gel or with size exclusion filtration columns before mixing them together to get rid of excess remainder strands.

X.Salt Buffers in Origami Patterning

Assumed Properties: Salt (Cation) Concentration, Thiol Gold Linkages

Affected System Goals: Task, Origami Patterning, Gold Patterning, Electrostatic Binding

Failure Type: Interaction

Description: Electrostatic binding approaches for the patterning of DNA origami require salt buffers with magnesium concentrations "incompatible with biochemical studies." Trying to reduce the salt buffers results in dissociation of DNA origami structures (Scheible et al. (2014)).

Origin: This problem motivated the invention of a protocol for patterning DNA origami without using electrostatic binding to avoid the salt buffer problem mentioned.

Cause: Salt concentrations are necessary in solution in order for DNA to hybridize. Other studies on electrostatic arrays have done more intense inquiries of this aspect (Fuchs et al. (2010)). Salts that bind with DNA often contain Magnesium2+ and Sodium+. The binding affects the polarity of the DNA's backbone, making salt an important factor in experiments. This effect has also been studied in Walter et al. (2013).

Detection: Evident from the design of the experiment.

Solution/Mitigation: origami can be attached using gold (which causes thiol-gold linkages with the origami) on passivated glass substrates that are etched using e-beam lithography. Samples were then measured using DNA-paint (Jungmann et al. (2014)) and TIRFM.

XI.Side Products and Thermal Degradation

Assumed Properties: Temperature, Annealing, Thermal ramp

Affected System Goals: Assembly, Yield, Time

Failure Type: Compound, Cumulative

Description: Origami (especially 3D) takes long to fold or forms along with many "molten-globule" side products when using a slow thermal ramping protocol (Sobczak et al. (2012)).

Origin: Sobczak et al. (2012) found that origami folds with higher yields at a set temperature that differs with each origami . They cite side products and thermal degradation as possible causes for low yields under a slow annealing protocol. The thermal ramp they used reduced the temperature by 1° C per every 3 hours.

Cause: A slow annealing protocol (also known as a thermal ramping protocol), created "molten globule" precursors that remained without turning into the intended product origami (Sobczak et al. (2012)).

Detection: Gel Electrophoresis can be used to determine if unintended side products are created due to the heating protocol. Sobczak et al. (2012) note that the precursors had greater "electrophoretic mobility", meaning that they can be identified by a distinct band on the gel.

Solution/Mitigation: Find the peak temperature at which the rate of folding is highest as described by Sobczak et al. (2012). Then set the temperature of the solution to the low boundary of the temperature peak. The solution is then heated to 65° C for 15 min (referred to as heat shock), then cooled immediately to the desired temperature. Yields were found to approach 100%, and folded within minutes instead of days.

XII.Unclear Structural Conformation

Assumed Properties: TEM, AFM, High Resolution, Cryo-TEM, Chirality, Symmetry, Static Structure

Affected System Goals: Analysis, Invasiveness

Failure Type: Component

Description: For detailed views of origami conformations the methods most popular have been atomic force microscopy, and transmission electron microscopy(TEM). Atomic Force Microscopy can alter the image, and has a lower resolution. All non-cryo TEM, though better, can also cause a structure to conform to a state that it would not in solution. Each method comes with some drawbacks, and neither arguably is state of the art.

Origin: A technique using cryo-electron microscopy, an extension of TEM, along with software based class averages was able to determine the shape of 3D origami with an overall resolution of 11.5 Angstroms (115 nm) (Bai et al. (2012)).

Cause: Viewing atoms at the scale of angstroms (1 A $= 1 * 10^{-10}$ meters) is challenging, albeit necessary for confirming the correct conformation of a device. Furthermore, alternative methods can further obscure a structural conformation by changing the structure of an origami. This method shown in Bai et al. (2012) offers one of the best ways to visually confirm the structure of origami while viewing it in solution.

Detection: When other viewing methods fail, this method should provide the most accuracy. However, it only works on chiral (asymmetrical), static (non-moving) structures.

Solution/Mitigation: Perform single particle cryo-electron microscopy as described by Bai et al. (2012). This will work best on chiral origami, as chirality is necessary for identifying the orientation of the scaffold. Bai et al. use liquid ethane to freeze the origami in solution before performing electron microscopy. The microscope can then scan for all of the origami. The images of origami can then be collated into 2D class averages and 3D reconstructions using software.

XIII.Unoptimized Staple Configuration

Assumed Properties: Seed Segment, Staple Domain, 3D origami

Affected System Goals: Assembly, Yield

Failure Type: Component, Cumulative

Description: Low yields of 3D origami are sometimes caused by inefficient staple placement. Specifically, staples lack a seed segment described below.

Origin: Ke et al. (2012) reported certain staple strategies that were unable to fold a 24 helix bundle (24-hb) rod. They then through experimentation found a method of stapling that increased the yield from effectively 0% to greater than 20% without changing the length of the staples.

Cause: Ke et al. hypothesize that for staple strands that lack a 14 nt seed the origami gets stuck kinetically in a misfolded intermediate stage. However, it is also possible that the problem is due to thermodynamics. (14 nt domains on average bind earlier and with more favorability than domains of lesser length as shown in SantaLucia (1998).) If it is due to kinetics, it is possible that the order in which staples bind to origami can be changed by altering the lengths of their domains binding with the scaffold.

Detection: If yield of origami is low (as shown by gel electrophoresis) or the shape is incorrect (as shown by electron microscopy of atomic force microscopy), the cause could be incorrect staple placement. As yields are higher for 2D origami, the staple configuration does not have as much of an effect.

Solution/Mitigation: Design staples such that they have a 14 nt seed domain. A seed domain is an uninterrupted stretch (no crossovers) of nt forming double stranded DNA with the scaffold. The seed domain is the part of the staple longer than the other domains, such that the seed domain binds first, hence it "seeds" the hybridization of the rest of the staple. It is also hypothesized that one can fine tune the times at which staples bind through these seed domains, leading to correct kinetic folding of the origami.

APPENDIX B. SUPPORTING INFORMATION FOR THE MODELING FRAMEWORK FOR DNA ORIGAMI WITH MULTIPLE CONFORMATIONS

B.1 More on the Model

The micro model does not include any information about the winding or orientation of the nucleotides (nt), and the macro model does not include information about winding, orientation, or stacking. Information is lost from the tertiary structure to the secondary structure of the micro model and from the micro model to the macro model. Physical systems do not necessarily correspond with either the tertiary, micro, or macro models. One benefit of using the macro model over the micro model or tertiary model is that the macro model does not make any implicit assumptions about which nucleotides are stacked or how the nucleotides are oriented.

To make the definition of the macro model unambiguous, let positions along the scaffold be labeled by the scaffold's phosphate backbones between the nucleotides. Let n_i, n_j be nucleotides on the scaffold. Let p_i, p_j be phosphates where p_i is adjacent to n_i and p_j is adjacent to n_j, such that neither p_i or p_j form a double stranded helix. (i.e. Any nucleotides bound to nucleotides next to p_i and p_j may not be connected by one backbone.) Finally, let s_{k_0}, s_{k_1} be nucleotides complementary to n_i, n_j respectively such that s_{k_0}, s_{k_1} are both adjacent to a phosphate backbone q_k. Nucleotide s_{k_0} will form hydrogen bonds with n_i and s_{k_1} will form hydrogen bonds with n_j. A **crossover** of degree 4 is then a set $\{p_i, p_j\}$ connected as follows:

$$p_i \leftrightarrow n_i \leftrightsquigarrow s_{k_0} \leftrightarrow q_k \leftrightarrow s_{k_1} \leftrightsquigarrow n_j \leftrightarrow p_j$$

A normal left-right arrow ($\leftrightarrow$) signifies a covalent bond, but does not say anything about the direction of the bond. A squiggly left-right ($\leftrightsquigarrow$) arrow signifies a hydrogen bond.

This definition of a crossover has a few notable features. First, it allows for coaxial helices to be parallel or antiparallel. Second, it also allows for crossovers of any degree $2n, n \in \mathbb{N}$ where n is the number of phosphate backbones being stapled together. Finally, it purposefully does not convey information about how the nucleotides in the crossover are oriented or stacked. This allows us to make equivalence classes over the macro-model, by ignoring differences in orientation or stacking.

B.2 Designing the Experiment

The structure I chose purposefully avoided all of the known mono-state motifs (Figure 2-5). The simplest such structure represented a square lattice of crossovers with two possible conformations. I made both conformations in CaDNAno (Douglas et al. (2009b)): the scaffold-stacked shown in Figure B.1 and the scaffold-stacked shown in Figure B.2. Due to its shape, I called this structure the rectangle.

Taking the staples generated from one design, I then added dumbbell hairpins(`dh = TCCT CTTT TGAG GAAC AAGT TTTC TTGT`) along the edges to differentiate one conformation from another (dumbbell sequences from Rothemund (2006)). The sequences can be found in Table B.1. The dumbbells along the edges are labeled in either green or red. The eleventh, yellow hairpin was added to give the origami chirality. The dumbbell insertion is colored in purple text. To verify that both of our structures used the same staple sequences, I generated staple sets for both of our structures and compared them (Table B.1).

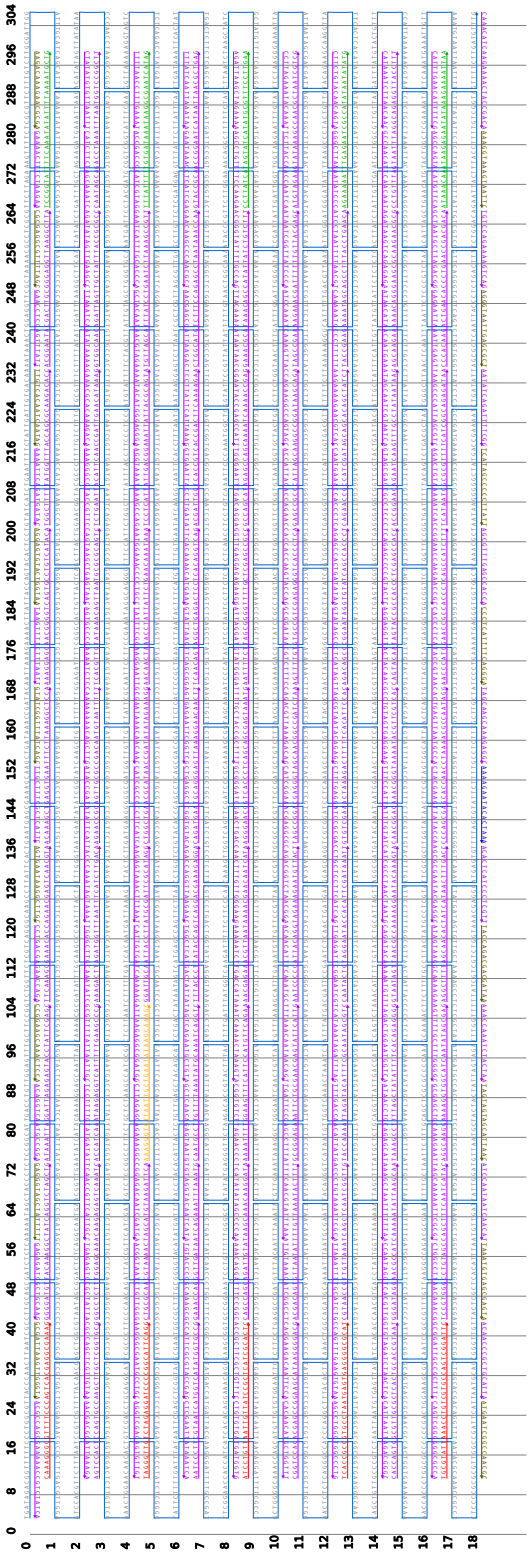


Figure B.1 Stacked Scaffold CaDNAno Design

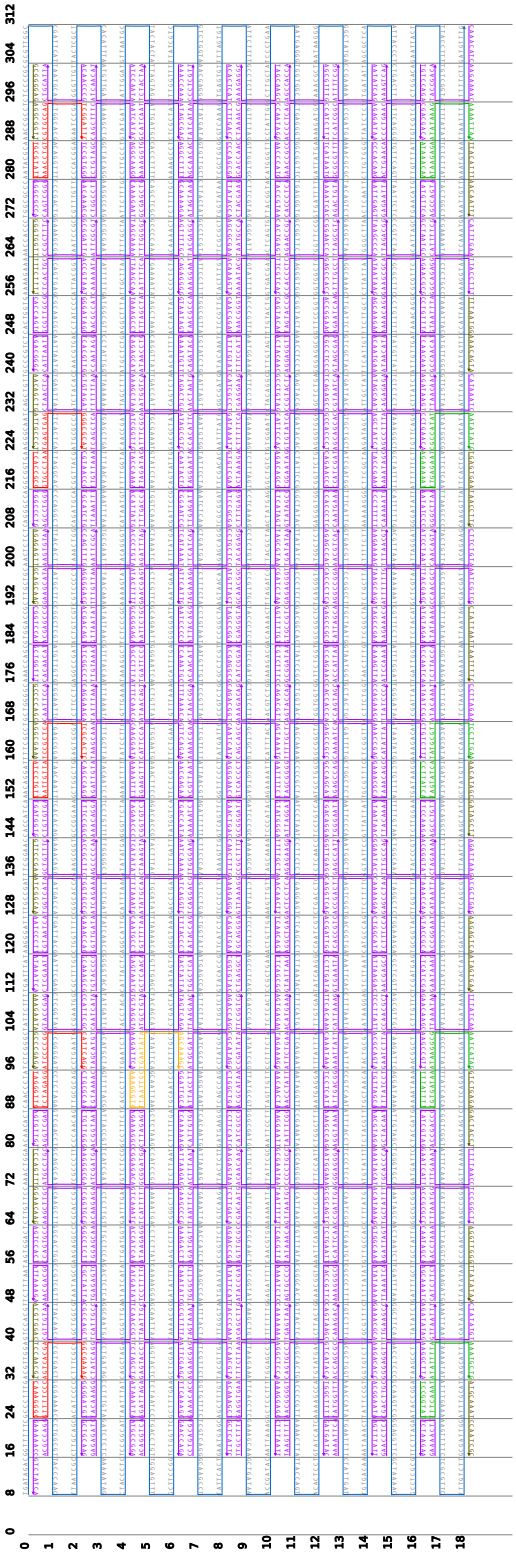


Figure B.2 Stacked Staple CaDNAno Design

Table B.1 Side by Side Identical Sequences for Both Conformations

Scaffold Stacked			Staple Stacked			
Start	End	Sequence	Start	End	Sequence	Length
8[122]	8[91]	AAAAATCTTGACCATAAATCAAAACTAAAGTA	7[144]	5[143]	AAAAATCTTGACCATAAATCAAAACTAAAGTA	32
9[75]	9[106]	AAAATTTTGGAAGTTTCATTCCATGTTCAGAA	4[159]	6[160]	AAAATTTTGGAAGTTTCATTCCATGTTCAGAA	32
8[58]	8[27]	AAACAGGATTTCCGGCACCGCTTCGCTGTTTC	3[144]	1[143]	AAACAGGATTTCCGGCACCGCTTCGCTGTTTC	32
3[107]	3[138]	AAAGACTTAGATGGTTTAATTTCATCTTGACA	6[63]	8[64]	AAAGACTTAGATGGTTTAATTTCATCTTGACA	32
1[171]	1[202]	AAAGGAGCAACGGGGTCAGTGCCTGTCATACA	10[31]	12[32]	AAAGGAGCAACGGGGTCAGTGCCTGTCATACA	32
2[218]	2[187]	AAAGGGCGTACCGTTCCAGTAAGCTGAGTAAC	13[48]	11[47]	AAAGGGCGTACCGTTCCAGTAAGCTGAGTAAC	32
5[75]	5[106]	AAAGGGTGTTAATTGCdhTGAATATAAGCAAAGC	4[95]	6[96]	AAAGGGTGTTAATTGCdhTGAATATAAGCAAAGC	32
5[171]	5[202]	AACAACTAAACCTATTATTCTGAACAAATAAA	10[95]	12[96]	AACAACTAAACCTATTATTCTGAACAAATAAA	32
17[171]	17[202]	AACCCATGGAACCGCCACCCTCAGCCATCTTT	10[287]	12[288]	AACCCATGGAACCGCCACCCTCAGCCATCTTT	32
2[26]	2[7]	AACGACGGATTAAAGAACGT	1[48]	0[44]	AACGACGGATTAAAGAACGT	20
9[107]	9[138]	AACGAGAAACGTTAATAAAACGAACAATCATA	6[159]	8[160]	AACGAGAAACGTTAATAAAACGAACAATCATA	32
6[250]	6[219]	AACGAGCGATTAAGACTCCTTATTCATTAAAG	15[112]	13[111]	AACGAGCGATTAAGACTCCTTATTCATTAAAG	32
11[107]	11[138]	AATCCCCCATTACAGGTAGAAAGAATGTTACT	6[191]	8[192]	AATCCCCCATTACAGGTAGAAAGAATGTTACT	32
9[171]	9[202]	AATTTTCTTAGCGGGGTTTTGCTCGAGCCGCC	10[159]	12[160]	AATTTTCTTAGCGGGGTTTTGCTCGAGCCGCC	32
1[139]	1[170]	ACAAAGCTGAGGTGAATTTCTTAAAGGCTCCA	8[31]	10[32]	ACAAAGCTGAGGTGAATTTCTTAAAGGCTCCA	32
10[122]	10[91]	ACAACATTTCAAATGCTTTAAACAATAACAGT	7[176]	5[175]	ACAACATTTCAAATGCTTTAAACAATAACAGT	32
3[75]	3[106]	ACCATCAAATAAGAGGTCATTTTTGAAGCCCG	4[63]	6[64]	ACCATCAAATAAGAGGTCATTTTTGAAGCCCG	32
10[218]	10[187]	ACCATTACACCAGAACCACCACCAAGTACCAG	13[176]	11[175]	ACCATTACACCAGAACCACCACCAAGTACCAG	32
4[282]	4[251]	ACCGACCGCATCGTAGGAATCATTTGCACCCA	17[80]	15[79]	ACCGACCGCATCGTAGGAATCATTTGCACCCA	32
15[203]	15[234]	ACCGGAACAGAATCAAGTTTGCCTATAATAAG	12[255]	14[256]	ACCGGAACAGAATCAAGTTTGCCTATAATAAG	32
4[302]	4[283]	ACCTTTTTAACCTTTGAAAT	18[83]	17[79]	ACCTTTTTAACCTTTGAAAT	20
11[43]	11[74]	ACGACAGTGTTAATATTTTGTTAAATACTTTT	2[191]	4[192]	ACGACAGTGTTAATATTTTGTTAAATACTTTT	32
7[203]	7[234]	ACGATTGGTCACCGTCACCGACTTCAAAAGAA	12[127]	14[128]	ACGATTGGTCACCGTCACCGACTTCAAAAGAA	32
18[14]	18[7]	ACGCGCGG	0[307]	0[300]	ACGCGCGG	8
1[235]	1[266]	ACGGAATACGACTTGCGGGAGGTTAAGGCTTA	14[31]	16[32]	ACGGAATACGACTTGCGGGAGGTTAAGGCTTA	32
4[58]	4[27]	ACGGTAATCTGTTGGGAAGGGCGAATGCCTGC	3[80]	1[79]	ACGGTAATCTGTTGGGAAGGGCGAATGCCTGC	32
1[7]	1[42]	ACGTCAAAGGGCGTTTTCCCdhAGTCACGAGGGCGAAAG	0[35]	2[32]	ACGTCAAAGGGCGTTTTCCCdhAGTCACGAGGGCGAAAG	36
13[267]	13[302]	AGAAAAATTTGAGAATdhCGCCATATAATATATGTGAG	16[223]	18[220]	AGAAAAATTTGAGAAdhTCGCCATATAATATATGTGAG	36
3[139]	3[170]	AGAACCGGAGTTGCGCCGACAATGTAATTTTT	8[63]	10[64]	AGAACCGGAGTTGCGCCGACAATGTAATTTTT	32
13[171]	13[202]	AGACAGCCCGGAATAGGTGTATCAGCCACCCT	10[223]	12[224]	AGACAGCCCGGAATAGGTGTATCAGCCACCCT	32
16[218]	16[187]	AGACTGTAAAAATCACCGGAACCAAGAACCGC	13[272]	11[271]	AGACTGTAAAAATCACCGGAACCAAGAACCGC	32
15[235]	15[266]	AGCAAGAAAAAACAGGGAAGCGCAAGAACGCG	14[255]	16[256]	AGCAAGAAAAAACAGGGAAGCGCAAGAACGCG	32
16[282]	16[251]	AGCCAGTACATGTTCAGCTAATGCTTAGACGG	17[272]	15[271]	AGCCAGTACATGTTCAGCTAATGCTTAGACGG	32
12[26]	12[7]	AGCCTGGGTGGCCCTGAGAG	1[208]	0[204]	AGCCTGGGTGGCCCTGAGAG	20
8[282]	8[251]	AGCCTGTTTTCCAAGAACGGGTATTTTGCCAG	17[144]	15[143]	AGCCTGTTTTCCAAGAACGGGTATTTTGCCAG	32
14[186]	14[155]	AGGAGGTTACTACAACGCCTGTAGTGAGGAAG	11[240]	9[239]	AGGAGGTTACTACAACGCCTGTAGTGAGGAAG	32
17[75]	17[106]	AGGCAAGGAAAAAGGTGGCATCAATATAGCGAG	4[287]	6[288]	AGGCAAGGAAAAAGGTGGCATCAATATAGCGAG	32
17[107]	17[138]	AGGCTTTTAGCAACACTATCATAACTTTGACC	6[287]	8[288]	AGGCTTTTAGCAACACTATCATAACTTTGACC	32
9[139]	9[170]	AGGGAACCGTCACCCTCAGCAGCGAGTAAATG	8[159]	10[160]	AGGGAACCGTCACCCTCAGCAGCGAGTAAATG	32
4[218]	4[187]	AGGTAAATAAAGCCAGAATGGAAACCCCTGCC	13[80]	11[79]	AGGTAAATAAAGCCAGAATGGAAACCCCTGCC	32
4[26]	4[7]	AGGTCGACAGTGTTGTTCCA	1[80]	0[76]	AGGTCGACAGTGTTGTTCCA	20
18[46]	18[23]	AGTAACAACCCGCTGCATTAATGA	2[307]	1[307]	AGTAACAACCCGCTGCATTAATGA	24
14[302]	14[283]	AGTACATAAATCTTAACAAC	18[243]	17[239]	AGTACATAAATCTTAACAAC	20
2[186]	2[155]	AGTGCCCGAAAATCTCCAAAAAAAACAGCTTG	11[48]	9[47]	AGTGCCCGAAAATCTCCAAAAAAAACAGCTTG	32
11[171]	11[202]	AGTTTTGTGTGCCGTCGAGAGGGTCACCCTCA	10[191]	12[192]	AGTTTTGTGTGCCGTCGAGAGGGTCACCCTCA	32
18[110]	18[87]	ATAAAAACCAAATCTACTAATAGT	6[307]	5[307]	ATAAAAACCAAATCTACTAATAGT	24
17[267]	17[302]	ATAAACAAATAAGAGAdhATATAAAGAAAATTAATTAC	16[287]	18[284]	ATAAACAAATAAGAGAdhATATAAAGAAAATTAATTAC	36
1[75]	1[106]	ATAAATTAGGATTAGAGAGTACCTATTCGAGC	4[31]	6[32]	ATAAATTAGGATTAGAGAGTACCTATTCGAGC	32
14[250]	14[219]	ATAACATAACAATGAAATAGCAATCGTAATCA	15[240]	13[239]	ATAACATAACAATGAAATAGCAATCGTAATCA	32
14[90]	14[59]	ATAACCTGCAGAGCATAAAGCTAATTTTTTAA	5[240]	3[239]	ATAACCTGCAGAGCATAAAGCTAATTTTTTAA	32
2[302]	2[283]	ATAACTATATGTGTTAATTT	18[51]	17[47]	ATAACTATATGTGTTAATTT	20
2[154]	2[123]	ATACCGATATATTCATTACCCAAATAGTAAAT	9[48]	7[47]	ATACCGATATATTCATTACCCAAATAGTAAAT	32
0[222]	0[199]	ATAGAAAATTCAGATGATACAGGA	13[12]	12[12]	ATAGAAAATTCAGATGATACAGGA	24
16[122]	16[91]	ATAGTAAGGCAAAAGAAGTTTTGCTTTGGGGC	7[272]	5[271]	ATAGTAAGGCAAAAGAAGTTTTGCTTTGGGGC	32
10[250]	10[219]	ATCCAAATCAAAGTTACCAGAAGGCCAGTAGC	15[176]	13[175]	ATCCAAATCAAAGTTACCAGAAGGCCAGTAGC	32
11[267]	11[302]	ATGTAGAAGTATAAAGCCAACGCTAATCGTCGCTAT	16[191]	18[188]	ATGTAGAAGTATAAAGCCAACGCTAATCGTCGCTAT	36
8[186]	8[155]	ATTAGGATGTATGGGATTTTGCTACCGCTTTT	11[144]	9[143]	ATTAGGATGTATGGGATTTTGCTACCGCTTTT	32
18[270]	18[247]	ATTCTGTCCAGAAAAGTCAGAGGG	16[307]	15[307]	ATTCTGTCCAGAAAAGTCAGAGGG	24
16[302]	16[283]	ATTTCATTTGAACATTTTCG	18[275]	17[271]	ATTTCATTTGAACATTTTCG	20
12[250]	12[219]	CAAAAATGCCCTTTTTAAGAAAAGAACGTCAC	15[208]	13[207]	CAAAAATGCCCTTTTTAAGAAAAGAACGTCAC	32
18[306]	18[279]	CAAACATCAAGAAAACTACCGACAAAAG	18[311]	17[307]	CAAACATCAAGAAAACTACCGACAAAAG	28
15[139]	15[170]	CAACGGAGAAACGGGTAAAATACGTTCGTCAC	8[255]	10[256]	CAACGGAGAAACGGGTAAAATACGTTCGTCAC	32
3[7]	3[42]	CAAGAGTCCACTCCAGTGCCAAGCTTGCTCGGTGCG	0[67]	2[64]	CAAGAGTCCACTCCAGTGCCAAGCTTGCTCGGTGCG	36
11[7]	11[42]	CAAGCGGTCCACACGAGCCGGAAGCATAAGGGGACG	0[195]	2[192]	CAAGCGGTCCACACGAGCCGGAAGCATAAGGGGACG	36
13[107]	13[138]	CAATACTGTAGGAATACCACATTCGATAAATT	6[223]	8[224]	CAATACTGTAGGAATACCACATTCGATAAATT	32
3[267]	3[302]	CAATAGCAGACCTAAATTTAATGGTCCGGCTTAGGT	16[63]	18[60]	CAATAGCAGACCTAAATTTAATGGTCCGGCTTAGGT	36
12[218]	12[187]	CAATGAAACCACCCTCAGAGCCACTGATATAA	13[208]	11[207]	CAATGAAACCACCCTCAGAGCCACTGATATAA	32
7[75]	7[106]	CAATGCCTTGTTTTAAATATGCAAATCAGGTC	4[127]	6[128]	CAATGCCTTGTTTTAAATATGCAAATCAGGTC	32
10[26]	10[7]	CACAACATGCTGGTTTGCCC	1[176]	0[172]	CACAACATGCTGGTTTGCCC	20
16[186]	16[155]	CACCCTCATACCGTAACACTGAGTTAATGCCA	11[272]	9[271]	CACCCTCATACCGTAACACTGAGTTAATGCCA	32
13[203]	13[234]	CAGAACCGCCATCGATAGCAGCACAGCTATCT	12[223]	14[224]	CAGAACCGCCATCGATAGCAGCACAGCTATCT	32
14[122]	14[91]	CAGATACAAAAATGTTTAGACTGGAATGGTCA	7[240]	5[239]	CAGATACAAAAATGTTTAGACTGGAATGGTCA	32
7[139]	7[170]	CAGATGAATGCAGGGAGTTAAAGGAACAACTT	8[127]	10[128]	CAGATGAATGCAGGGAGTTAAAGGAACAACTT	32
9[43]	9[74]	CAGCCAGCAGATTGTATAAGCAAACAAGGATA	2[159]	4[160]	CAGCCAGCAGATTGTATAAGCAAACAAGGATA	32
0[158]	0[135]	CAGCTTGCTTTCGCTCATTCAGTG	9[12]	8[12]	CAGCTTGCTTTCGCTCATTCAGTG	24
15[171]	15[202]	CAGTACAATAGTACCGCCACCCTCGAGCCACC	10[255]	12[256]	CAGTACAATAGTACCGCCACCCTCGAGCCACC	32
16[26]	16[7]	CAGTCGGGGGGCGCCAGGGT	1[272]	0[268]	CAGTCGGGGGGCGCCAGGGT	20
4[154]	4[123]	CATCGCCCCCTTCATCAAGAGTAAACTTTAAT	9[80]	7[79]	CATCGCCCCCTTCATCAAGAGTAAACTTTAAT	32
10[154]	10[123]	CATCGGAAACGAGGCGCAGACGGTCTAACGGA	9[176]	7[175]	CATCGGAAACGAGGCGCAGACGGTCTAACGGA	32
2[282]	2[251]	CATCTTCTAGCAAATCAGATATAGTTGAAGCC	17[48]	15[47]	CATCTTCTAGCAAATCAGATATAGTTGAAGCC	32
14[26]	14[7]	CATTAATTAGACGGGCAACA	1[240]	0[236]	CATTAATTAGACGGGCAACA	20
4[122]	4[91]	CATTGTGATCAAAAAGATTAAGAGGCGGATGG	7[80]	5[79]	CATTGTGATCAAAAAGATTAAGAGGCGGATGG	32
14[58]	14[27]	CCAATAGGTCACGTTGGTGTAGATCTAACTCA	3[240]	1[239]	CCAATAGGTCACGTTGGTGTAGATCTAACTCA	32
17[235]	17[266]	CCCACAAGACTGAACACCCTGAACCGACGACA	14[287]	16[288]	CCCACAAGACTGAACACCCTGAACCGACGACA	32
17[139]	17[170]	CCCAGCGAGCACCAACCTAAAACGCCAATAGG	8[287]	10[288]	CCCAGCGAGCACCAACCTAAAACGCCAATAGG	32
13[7]	13[42]	CCCTTCACCGCCGTGCCTAAdhTGAGTGAGGGGCGCAT	0[227]	2[224]	CCCTTCACCGCCGTGCCTAAdhTGAGTGAGGGGCGCAT	36
15[267]	15[302]	CCTGTTTAGTAATTTAGGCAGAGGTTACCTTTTTTA	16[255]	18[252]	CCTGTTTAGTAATTTAGGCAGAGGTTACCTTTTTTA	36
9[267]	9[302]	CCTTATCATAGTATCAdhTATGCGTTTCCTTGAAAACA	16[159]	18[156]	CCTTATCATAGTATCAdhTATGCGTTTCCTTGAAAACA	36
18[238]	18[215]	CGCTAATATCAGATTTTCGGTCAT	14[307]	13[307]	CGCTAATATCAGATTTTCGGTCAT	24
8[90]	8[59]	CGGTGTCTTAGAACCCTCATATATAGCCCCAA	5[144]	3[143]	CGGTGTCTTAGAACCCTCATATATAGCCCCAA	32

Table B.1 continued

13[43]	13[74]	CGTAACCGTTGTTAAATCAGCTCAATCGGTTG	2[223]	4[224]	CGTAACCGTTGTTAAATCAGCTCAATCGGTTG	32
16[154]	16[123]	CTACGAAGTTATACCAAGCGCGAATACGAGGC	9[272]	7[271]	CTACGAAGTTATACCAAGCGCGAATACGAGGC	32
6[26]	6[7]	CTCGAATTATAAATCAAAAG	1[112]	0[108]	CTCGAATTATAAATCAAAAG	20
6[154]	6[123]	CTGAGGCTCGGTGTACAGACCAGGTTAAGAAC	9[112]	7[111]	CTGAGGCTCGGTGTACAGACCAGGTTAAGAAC	32
5[43]	5[74]	CTGCGCAACGTAAAACTAGCATGTAAGATTCA	2[95]	4[96]	CTGCGCAACGTAAAACTAGCATGTAAGATTCA	32
7[235]	7[266]	CTGGCATGTCTTTCCAGAGCCTAATAAACCAA	14[127]	16[128]	CTGGCATGTCTTTCCAGAGCCTAATAAACCAA	32
16[58]	16[27]	CTGGCCTTAACAAACGGCGGATTGCCGCTTTC	3[272]	1[271]	CTGGCCTTAACAAACGGCGGATTGCCGCTTTC	32
5[139]	5[170]	CTGGCTGAACGCATAACCGATATATAGAAAGG	8[95]	10[96]	CTGGCTGAACGCATAACCGATATATAGAAAGG	32
8[26]	8[7]	CTGTGTGATGATGGTGGTTC	1[144]	0[140]	CTGTGTGATGATGGTGGTTC	20
4[90]	4[59]	CTTAGAGCAGAAAGGCCGGAGACAATCGATGA	5[80]	3[79]	CTTAGAGCAGAAAGGCCGGAGACAATCGATGA	32
12[302]	12[283]	CTTGCTTCTGTACAACAGTA	18[211]	17[207]	CTTGCTTCTGTACAACAGTA	20
12[154]	12[123]	CTTTGAGGATCCGCGACCTGCTCCTTCATCAG	9[208]	7[207]	CTTTGAGGATCCGCGACCTGCTCCTTCATCAG	32
9[7]	9[42]	GAAAATCCTGTTAATTGTTAATTCCGCTCATCGCACTC	0[163]	2[160]	GAAAATCCTGTTAATTGTTAATTCCGCTCATCGCACTC	36
7[43]	7[74]	GAAACCAGGGTTGATAATCAGAAATTTAAATG	2[127]	4[128]	GAAACCAGGGTTGATAATCAGAAATTTAAATG	32
9[235]	9[266]	GAAACGCATAAACAGCCATATTATCTGTCTTT	14[159]	16[160]	GAAACGCATAAACAGCCATATTATCTGTCTTT	32
6[282]	6[251]	GAATAAACCTCATCGAGAACAAGCCCAACGCT	17[112]	15[111]	GAATAAACCTCATCGAGAACAAGCCCAACGCT	32
16[250]	16[219]	GAGAATTAAATTGAGTTAAGCCCATTAGCGTC	15[272]	13[271]	GAGAATTAAATTGAGTTAAGCCCATTAGCGTC	32
5[7]	5[42]	GAGATAGGGTTGTCTAGAGGATCCCCGGCATTCAGG	0[99]	2[96]	GAGATAGGGTTGTCTAGAGGATCCCCGGCATTCAGG	36
11[203]	11[234]	GAGCCGCCCATTAGCAAGGCCGGATAAGCAGA	12[191]	14[192]	GAGCCGCCCATTAGCAAGGCCGGATAAGCAGA	32
7[7]	7[42]	GCAAAATCCCTTCGTAATCATGGTCATATGGTGCCG	0[131]	2[128]	GCAAAATCCCTTCGTAATCATGGTCATATGGTGCCG	36
14[282]	14[251]	GCCAACATTCAACAATAGATAAGTCAGAGAGA	17[240]	15[239]	GCCAACATTCAACAATAGATAAGTCAGAGAGA	32
9[203]	9[234]	GCCAGCATAGAGCCAGCAAAATCAAAACCGAG	12[159]	14[160]	GCCAGCATAGAGCCAGCAAAATCAAAACCGAG	32
0[286]	0[263]	GCGAGAAAACTTTCTAAGAACGCG	17[12]	16[12]	GCGAGAAAACTTTCTAAGAACGCG	24
16[90]	16[59]	GCGAGCTGCAAAGAATTAGCAAAAATTCGCGT	5[272]	3[271]	GCGAGCTGCAAAGAATTAGCAAAAATTCGCGT	32
10[186]	10[155]	GCGGATAACGTCTTTCCAGACGTTAAAGACAG	11[176]	9[175]	GCGGATAACGTCTTTCCAGACGTTAAAGACAG	32
11[75]	11[106]	GCGGGAGAAATTCTGCGAACGAGTATTCATTG	4[191]	6[192]	GCGGGAGAAATTCTGCGAACGAGTATTCATTG	32
8[154]	8[123]	GCGGGATCGAACTGACCAACTTTGTTGGGAAG	9[144]	7[143]	GCGGGATCGAACTGACCAACTTTGTTGGGAAG	32
4[250]	4[219]	GCTACAATAACGTAGAAAATACATGGGAGGGA	15[80]	13[79]	GCTACAATAACGTAGAAAATACATGGGAGGGA	32
6[90]	6[59]	GCTCAACAGAGTAATGTGTAGGTACAATCATA	5[112]	3[111]	GCTCAACAGAGTAATGTGTAGGTACAATCATA	32
8[302]	8[283]	GCTTAGATTAAGCTAGAAAA	18[147]	17[143]	GCTTAGATTAAGCTAGAAAA	20
5[107]	5[138]	GGATTGCAATTACCTTATGCGATTCGCATAGG	6[95]	8[96]	GGATTGCAATTACCTTATGCGATTCGCATAGG	32
3[43]	3[74]	GGCCTCTTTGGAGCAAACAAGAGAGTCAAATC	2[63]	4[64]	GGCCTCTTTGGAGCAAACAAGAGAGTCAAATC	32
18[174]	18[151]	GGGATAGCAAGCAAAGAGGCAAAA	10[307]	9[307]	GGGATAGCAAGCAAAGAGGCAAAA	24
15[43]	15[74]	GGGATAGGAACGCCATCAAAAATATTAAGCAA	2[255]	4[256]	GGGATAGGAACGCCATCAAAAATATTAAGCAA	32
12[282]	12[251]	GGGCTTAAAATATCCCATCCTAATTTTAACGT	17[208]	15[207]	GGGCTTAAAATATCCCATCCTAATTTTAACGT	32
1[43]	1[74]	GGGGATGTTACAAAGGCTATCAGGTCTAGCTG	2[31]	4[32]	GGGGATGTTACAAAGGCTATCAGGTCTAGCTG	32
0[30]	0[3]	GGTAACGCCAGGGAAAAACCGTCTATCA	1[12]	0[8]	GGTAACGCCAGGGAAAAACCGTCTATCA	28
17[7]	17[42]	GGTTTGCGTATTAAACCTGTTACGTGCCAGTCGGATTC	0[291]	2[288]	GGTTTGCGTATTAAACCTGTTACGTGCCAGTCGGATTC	36
0[190]	0[167]	GTAATAAGTTTTCTTTAATTGTAT	11[12]	10[12]	GTAATAAGTTTTCTTTAATTGTAT	24
15[107]	15[138]	GTAATAGTTAACGCCAAAAGGAATACAAAGTA	6[255]	8[256]	GTAATAGTTAACGCCAAAAGGAATACAAAGTA	32
7[267]	7[302]	GTACCGCAACCGGAATCATAATTAACGCTGAGAAGA	16[127]	18[124]	GTACCGCAACCGGAATCATAATTAACGCTGAGAAGA	36
14[218]	14[187]	GTAGCGACCGCCTCCCTCAGAGCCCCGTACTC	13[240]	11[239]	GTAGCGACCGCCTCCCTCAGAGCCCCGTACTC	32
12[186]	12[155]	GTATAGCCCTCATAGTTAGCGTAATACAGAGG	11[208]	9[207]	GTATAGCCCTCATAGTTAGCGTAATACAGAGG	32
6[186]	6[155]	GTATTAAGTTCAGCGGAGTGAGAATTCGGTCG	11[112]	9[111]	GTATTAAGTTCAGCGGAGTGAGAATTCGGTCG	32
6[218]	6[187]	GTGAATTACCTTGATATTCACAAAACATGAAA	13[112]	11[111]	GTGAATTACCTTGATATTCACAAAACATGAAA	32
3[235]	3[266]	GTGGCAACAGATTAGTTGCTATTTACCGCGCC	14[63]	16[64]	GTGGCAACAGATTAGTTGCTATTTACCGCGCC	32
13[139]	13[170]	GTGTCGAAACTAAAGACTTTTTCACATTCCAC	8[223]	10[224]	GTGTCGAAACTAAAGACTTTTTCACATTCCAC	32
18[142]	18[119]	TAAAACACTCATCCCTCGTTTACC	8[307]	7[307]	TAAAACACTCATCCCTCGTTTACC	24
15[75]	15[106]	TAAAGCCTTTTAGCTATATTTTCACAGAGGGG	4[255]	6[256]	TAAAGCCTTTTAGCTATATTTTCACAGAGGGG	32
18[78]	18[55]	TAACATCCAATAATCAACATTAAA	4[307]	3[307]	TAACATCCAATAATCAACATTAAA	24
13[75]	13[106]	TACCAAAATTAGATACATTTCGCAATAGCGTC	4[223]	6[224]	TACCAAAATTAGATACATTTCGCAATAGCGTC	32
13[235]	13[266]	TACCGAAGAAAATAGCAGCCTTTACCTGAACA	14[223]	16[224]	TACCGAAGAAAATAGCAGCCTTTACCTGAACA	32
11[235]	11[266]	TAGCCGAAAAGAAACGATTTTTTGTTACGAGC	14[191]	16[192]	TAGCCGAAAAGAAACGATTTTTTGTTACGAGC	32
11[139]	11[170]	TAGCCGGACGAGGGTAGCAACGGCCGATCTAA	8[191]	10[192]	TAGCCGGACGAGGGTAGCAACGGCCGATCTAA	32
0[254]	0[231]	TAGCGAACCTCCAGTTTATTTTGT	15[12]	14[12]	TAGCGAACCTCCAGTTTATTTTGT	24
18[206]	18[183]	TATTAGCGTTTGAGCCACCACCCT	12[307]	11[307]	TATTAGCGTTTGAGCCACCACCCT	24
4[186]	4[155]	TATTTCGGAAGGAATTGCGAATAAACAACAAC	11[80]	9[79]	TATTTCGGAAGGAATTGCGAATAAACAACAAC	32
7[171]	7[202]	TCAACAGTAGGCTGAGACTCCTCACAGGTCAG	10[127]	12[128]	TCAACAGTAGGCTGAGACTCCTCACAGGTCAG	32
3[171]	3[202]	TCACGTTGTATAAACAGTTAATGCGCGCAGTC	10[63]	12[64]	TCACGTTGTATAAACAGTTAATGCGCGCAGTC	32
17[203]	17[234]	TCATAATCGCGCGTTTTCATCGGCAGAGATAA	12[287]	14[288]	TCATAATCGCGCGTTTTCATCGGCAGAGATAA	32
0[94]	0[71]	TCCAACAGGTCAATGCCGGAGAGG	5[12]	4[12]	TCCAACAGGTCAATGCCGGAGAGG	24
1[267]	1[302]	TCCGGTATTTTCAAATGGATATTTTAAAATGCTGATGC	16[31]	18[28]	TCCGGTATTTTCAAATGGATATTTTAAAATGCTGATGC	36
17[43]	17[74]	TCCGTGGGCCTGTAGCCAGCTTTCAATCATAC	2[287]	4[288]	TCCGTGGGCCTGTAGCCAGCTTTCAATCATAC	32
5[203]	5[234]	TCCTCATTATTGACGGAAATTATTACGCAGTA	12[95]	14[96]	TCCTCATTATTGACGGAAATTATTACGCAGTA	32
2[90]	2[59]	TCCTTTTGTATGATATTCAACCGTTCATTGCC	5[48]	3[47]	TCCTTTTGTATGATATTCAACCGTTCATTGCC	32
0[302]	0[295]	TCGCAAGA	18[19]	18[12]	TCGCAAGA	8
3[203]	3[234]	TCTGAATTACATTCAACCGATTGAACATAAAG	12[63]	14[64]	TCTGAATTACATTCAACCGATTGAACATAAAG	32
10[282]	10[251]	TCTTACCAACCAATCAATAATCGGTTATCCCA	17[176]	15[175]	TCTTACCAACCAATCAATAATCGGTTATCCCA	32
6[302]	6[283]	TGAATTTATCAATTAAATAA	18[115]	17[111]	TGAATTTATCAATTAAATAA	20
2[58]	2[27]	TGAGAGTCCGCTATTACGCCAGCTCGTTGTAA	3[48]	1[47]	TGAGAGTCCGCTATTACGCCAGCTCGTTGTAA	32
10[90]	10[59]	TGATTCCCAGCCTTTATTTCAACGTATTTAAA	5[176]	3[175]	TGATTCCCAGCCTTTATTTCAACGTATTTAAA	32
6[122]	6[91]	TGGCTCATGACTATTATAGTCAGAATGCTGTA	7[112]	5[111]	TGGCTCATGACTATTATAGTCAGAATGCTGTA	32
1[203]	1[234]	TGGCTTTTTATGGTTTACCAGCGCAAGACACC	12[31]	14[32]	TGGCTTTTTATGGTTTACCAGCGCAAGACACC	32
8[218]	8[187]	TGGGAATTTGACAGGAGGTTGAGGAGAGAAGG	13[144]	11[143]	TGGGAATTTGACAGGAGGTTGAGGAGAGAAGG	32
2[122]	2[91]	TGGGCTTGCAAATATCGCGTTTTATTAATTGC	7[48]	5[47]	TGGGCTTGCAAATATCGCGTTTTATTAATTGC	32
6[58]	6[27]	TGTACCCCGCAAAGCGCCATTCGCGTACCGAG	3[112]	1[111]	TGTACCCCGCAAAGCGCCATTCGCGTACCGAG	32
5[235]	5[266]	TGTTAGCATTTATCCTGAATCTTAAAGCCGTT	14[95]	16[96]	TGTTAGCATTTATCCTGAATCTTAAAGCCGTT	32
2[250]	2[219]	TTAAATCAATATAAAAGAAACGCACAAAGACA	15[48]	13[47]	TTAAATCAATATAAAAGAAACGCACAAAGACA	32
12[58]	12[27]	TTAAATTTTGCATCTGCCAGTTTGAAGTGTAA	3[208]	1[207]	TTAAATTTTGCATCTGCCAGTTTGAAGTGTAA	32
8[250]	8[219]	TTACAAAAATAATAACGGAATACCGAGCCATT	15[144]	13[143]	TTACAAAAATAATAACGGAATACCGAGCCATT	32
1[107]	1[138]	TTCAAAGCGAAACACCAGAACGAGTCAACGTA	6[31]	8[32]	TTCAAAGCGAAACACCAGAACGAGTCAACGTA	32
12[122]	12[91]	TTGAGATTCGGAATCGTCATAAATAGATTTAG	7[208]	5[207]	TTGAGATTCGGAATCGTCATAAATAGATTTAG	32
0[126]	0[103]	TTGCCCTGACGAGAACCAGACCGG	7[12]	6[12]	TTGCCCTGACGAGAACCAGACCGG	24
10[58]	10[27]	TTGTAAACATCGGCCTCAGGAAGACAATTCCA	3[176]	1[175]	TTGTAAACATCGGCCTCAGGAAGACAATTCCA	32
7[107]	7[138]	TTTACCCTTATACCAGTCAGGACGAAAGAGGA	6[127]	8[128]	TTTACCCTTATACCAGTCAGGACGAAAGAGGA	32
5[267]	5[302]	TTTATTTTTGTGATAAAGATAAGGCGAATCATAGGTCT	16[95]	18[92]	TTTATTTTTGTGATAAAGATAAGGCGAATCATAGGTCT	36
14[154]	14[123]	TTTCCATTATTTGTATCATCGCCTAACTAATG	9[240]	7[239]	TTTCCATTATTTGTATCATCGCCTAACTAATG	32
10[302]	10[283]	TTTCCCTTAGAAATACAAAT	18[179]	17[175]	TTTCCCTTAGAAATACAAAT	20
12[90]	12[59]	TTTGACCAACATTATGACCCTGTAAATTCGCA	5[208]	3[207]	TTTGACCAACATTATGACCCTGTAAATTCGCA	32
15[7]	15[42]	TTTTCACCAGTGGCGTTGCGCTCACTGCACCGTAAT	0[259]	2[256]	TTTTCACCAGTGGCGTTGCGCTCACTGCACCGTAAT	36
0[62]	0[39]	TTTTGAGAGATCGCTGCAAGGCGA	3[12]	2[12]	TTTTGAGAGATCGCTGCAAGGCGA	24

B.3 Estimating Free Energy

To calculate the difference in free energy between the structures, I wrote a program to count the occurrence of different quadnucleotides in each crossover (nucleotides *abcd* in Figure 1*E*). The results of this count for an m13mp18 scaffold are shown in Data B.1 (*abcd* left and $a'b'c'd'$ right). I then used this count to find the number of coaxially stacked nearest neighbors. I calculated the entropy, enthalpy, and Gibb's free energy using nearest neighbor thermodynamic parameters (Peyret (2000) and SantaLucia Jr and Hicks (2004)). The resultant free energies for both were calculated vs. temperature (Data B.2). The energy landscape in Figure 7*E* was made using the values at 323K.

Data B.1 Coaxial Stacked Nucleotides Count

Scaffold Stacked				Staple Stacked			
AAAA	1			AAAA	4		
AAAC	3			AAAC	3		
AACA	4			AAAG	5		
AACG	3	CTTT	5	AACA	4	CGTT	
AATA	2	GAAC	1	AACC	4	CTAA	
AATG	1	GAAG	1	AAGA	3	CTAC	
AATT	8	GACA	1	AAGC	3	CTCA	
ACCG	1	GACT	2	AAGT	3	CTGA	
ACCT	3	GATC	1	AATA	6	CTGG	
ACGT	1	GCAA	1	AATC	3	CTTA	
ACTA	3	GCAC	1	AATG	2	CTTG	
ACTG	1	GCGC	1	AATT	4	GAGC	
ACTT	1	GCTG	1	ACCA	2	GATT	
AGAT	2	GGAA	1	ACCC	2	GCCG	
AGCT	2	GGAC	1	ACCG	2	GCCT	
AGTG	4	GGCC	1	ACGT	2	GCGC	
ATAA	2	GGGG	1	ACTG	2	GCGG	
ATAT	2	GGGT	2	AGAC	1	GCGT	
ATCT	3	GGTA	1	AGCA	2	GCTA	
ATGA	1	GGTT	2	AGCG	3	GCTT	
ATGC	2	GTAA	1	AGCT	1	GGGT	
ATGG	2	GTCA	1	AGTA	2	GTTT	
ATTA	4	GTTA	3	AGTG	4	TAAC	
ATTC	3	GTTG	4	ATAG	3	TACC	
ATTG	1	GTTT	3	ATAT	2	TAGG	
CAAG	2	TACA	1	ATCG	2	TATA	
CACG	1	TATA	1	ATCT	2	TATG	
CACT	4	TATG	1	ATGA	1	TCAG	
CATT	2	TATT	6	ATGC	1	TCAT	
CCAC	1	TCAG	1	ATGG	3	TCCA	
CCGA	1	TCCA	1	ATTA	8	TCCT	
CCGC	2	TCGC	1	CAAT	2	TCGT	
CCTA	1	TCGG	1	CACG	1	TCTG	
CGAT	1	TCGT	3	CAGC	1	TGAT	
CGCT	3	TCTT	3	CAGT	1	TGGA	
CGTA	4	TGCG	1	CATA	1	TGGC	
CGTT	2	TGGA	1	CATG	4	TTAC	
CTAA	3	TGTT	4	CCCC	1	TTAT	
CTGG	2	TTAG	2	CCGA	1	TTTA	
CTGT	1	TTGA	1	CCTT	1	TTTG	
CTTA	2	TTTT	4	CGAA	2	TTTT	
CTTG	2			CGTG	1		

Data B.2 Calculation of Free Energy

			Scaffold stacked pairs			Staple stacked pairs		
NN	NN stack H(kcal/mol)	NN stack S(cal/molK)	Quantity	H(kcal/mol)	S(kcal/mol)	Quantity	H(kcal/mol)	S(kcal/mol)
AA	-9.9	-23.1	25	-247.5	-0.5775	47	-465.3	-1.0857
AC	-12	-33.1	12	-144	-0.3972	19	-228	-0.6289
AG	-13.5	-34.1	18	-243	-0.6138	32	-432	-1.0912
AT	-9.7	-26.2	32	-310.4	-0.8384	25	-242.5	-0.655
CA	-20.2	-55.5	27	-545.4	-1.4985	21	-424.2	-1.1655
CC	-8	-20.6	8	-64	-0.1648	15	-120	-0.309
CG	-13.1	-31.4	14	-183.4	-0.4396	16	-209.6	-0.5024
CT	-17.7	-48.5	22	-389.4	-1.067	14	-247.8	-0.679
GA	-12.8	-33.5	9	-115.2	-0.3015	14	-179.2	-0.469
GC	-9	-26.4	16	-144	-0.4224	10	-90	-0.264
GG	-18.7	-52	11	-205.7	-0.572	6	-112.2	-0.312
GT	-16.2	-16.8	18	-291.6	-0.3024	17	-275.4	-0.2856
TA	-20.5	-59.2	27	-553.5	-1.5984	30	-615	-1.776
TC	-9	-25.4	11	-99	-0.2794	10	-90	-0.254
TG	-10.5	-27.8	23	-241.5	-0.6394	17	-178.5	-0.4726
TT	-15.3	-44.7	51	-780.3	-2.2797	31	-474.3	-1.3857
		Total	324	-4557.9	-11.992	324	-4384	-11.3356

Total Crossovers = 162

Temp (K)	Scaffold G	Staple G	Normalized Diff.
273	-1284.084	-1289.3812	0.004125275
283	-1164.164	-1176.0252	0.010188599
293	-1044.244	-1062.6692	0.017644535
303	-924.324	-949.3132	0.027035109
313	-804.404	-835.9572	0.039225563
323	-684.484	-722.6012	0.055687496
333	-564.564	-609.2452	0.079142843
343	-444.644	-495.8892	0.115249953
353	-324.724	-382.5332	0.178025646
363	-204.804	-269.1772	0.314316127
373	-84.884	-155.8212	0.835695773

Table B.2 Influence Strands for Staple and Scaffold Stacked Conformations

Staple Stacked			Scaffold Stacked		
Start	End	Sequence	Start	End	Sequence
18[27]	18[20]	AAATCCAA	0[102]	0[95]	AAGCAAAC
0[107]	0[100]	AATAGCCC	0[134]	0[127]	AATAAGGC
0[203]	0[196]	AGTTGCAG	18[118]	18[111]	AGACGACG
18[251]	18[244]	ATGGAAAC	18[214]	18[207]	AGCCCCCT
18[283]	18[276]	ATTTAACA	0[262]	0[255]	AGGCGTTT
0[171]	0[164]	CAGCAGGC	18[86]	18[79]	AGTAGCAT
0[139]	0[132]	CGAAATCG	18[22]	18[15]	ATCGGCCA
18[91]	18[84]	GAGAGACT	0[294]	0[287]	CAAAGAAC
0[235]	0[228]	GCTGATTG	0[230]	0[223]	CACAATCA
0[43]	0[36]	GGACTCCA	18[182]	18[175]	CATTTTCA
0[299]	0[292]	GGAGAGGC	0[166]	0[159]	CGGTTTAT
0[267]	0[260]	GGTTTTTC	18[150]	18[143]	GAATACAC
18[123]	18[116]	GTCAATAG	18[278]	18[271]	GTAAAGTA
0[75]	0[68]	GTTTGGAA	0[70]	0[63]	GTAGCTAT
18[187]	18[180]	TAATTAAT	0[198]	0[191]	GTGTACTG
18[155]	18[148]	TAGCGATA	18[246]	18[239]	TAATTGAG
18[219]	18[212]	TGAATAAC	18[54]	18[47]	TGTGAGCG
18[59]	18[52]	TGGGTTAT	0[38]	0[31]	TTAAGTTG

The influencer strands in Table B.2 hybridize with the single stranded regions of the rectangle (shown as dashed edges in Figure 7*A*). Both conformations have 18 influencer domains, thus each conformation has 18 16nt long influencer strands.

B.4 Observations

To count the structures we used fluid Tapping Mode^{TM} on a Digital Instruments MultiMode Atomic Force Microscopy (AFM) with AC40 AFM probes from Bruker. Figure B.3 shows the output from the AFM. The rectangles and dumbbell hairpins were visible due to their height on the mica. Researchers were given a key for how to interpret results shown in Figure B.4. In addition to keeping track of whether the origami was face up or face down, researchers were also asked to keep track of the orientation (relative to the horizon with east = 0° and north = 90°). Table B.3 is an example count for Figure B.3.

Three researchers independently (blind) counted each AFM output. Their results were compared in order to create an estimate of the average and standard error for each experiment in Table B.3. Each count was treated as a Bernoulli distribution due to the binary nature of the sample (each origami was either staple-stacked or scaffold-stacked with high probability,

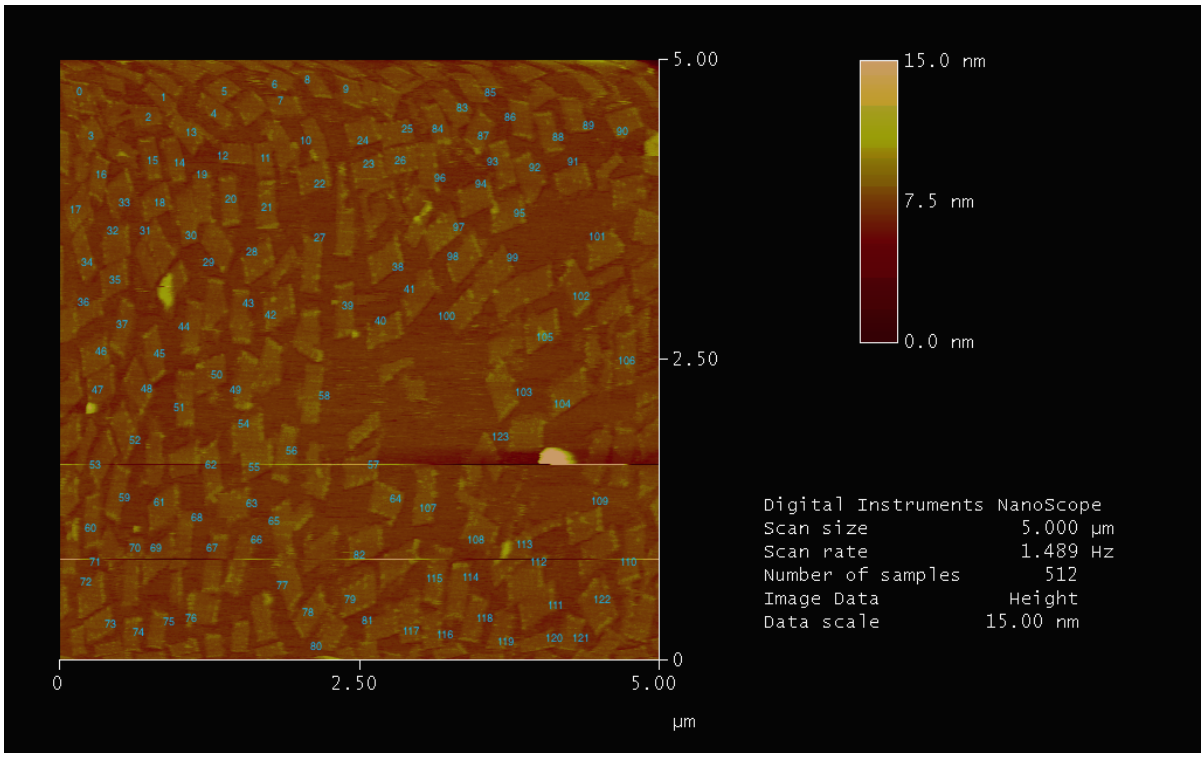


Figure B.3 Example AFM with Numbered Rectangles

and for the No Influencer experiment 92.7% of rectangles had visible hairpins). The counts were compared to derive the Standard error, which we used for our error bars in Figure 9*B*. For certain experiments, not enough structures were counted to derive statistics, in which case the column contains "NA."

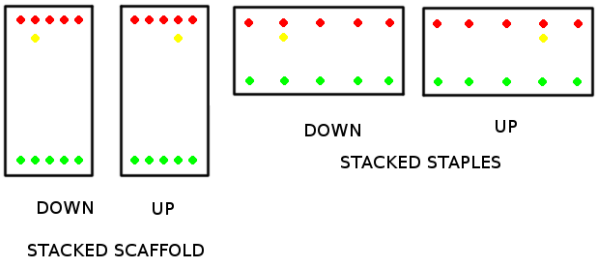


Figure B.4 Key for Stacked Staple/Scaffold, Up/Down

Table B.3 Example Counting Sheet

Origami #	Rectangle? (Y/N)	(U)p (D)own or (O)ther	s(C)af s(T)ap or (O)ther	Rotation [-60, 90]
0	Y	O	C	-60
1	Y	O	C	-60
2	Y	O	T	0
3	Y	O	T	30
4	Y	U	C	-30
5	Y	O	T	30
6	Y	O	O	30
7	Y	O	C	0
8	Y	O	C	90
9	Y	D	C	90
10	Y	O	O	-60
11	Y	O	C	90
12	Y	O	T	0
13	Y	O	C	-30
14	Y	O	C	90
15	Y	D	C	90
16	Y	D	C	90
17	Y	O	T	60
18	Y	U	C	30
19	Y	O	T	90
20	Y	O	C	90
21	Y	O	C	90
22	Y	D	C	30
23	Y	O	C	30
24	Y	O	C	30
25	Y	U	T	30
26	Y	U	T	30
27	Y	O	T	90
28	Y	U	C	30
29	Y	O	T	-30
30	Y	O	O	90
31	Y	O	T	-60
32	Y	O	C	0
33	Y	O	C	30
34	Y	O	T	-60
35	Y	O	C	30
36	Y	O	C	30
37	Y	D	C	60
38	Y	D	C	30
39	Y	O	C	90
40	Y	O	C	0
41	Y	D	C	60
42	Y	U	C	30
43	Y	D	T	90
44	Y	O	C	90
45	Y	O	C	-60
46	Y	U	C	30
47	Y	D	C	30
48	Y	D	C	-30
49	Y	O	C	60
50	Y	O	T	30
51	Y	D	C	90
52	Y	U	T	-30
53	Y	D	C	60
54	Y	D	C	30
55	Y	O	C	30
56	Y	O	C	60
57	Y	O	T	30
58	Y	O	O	90
59	Y	O	C	60
60	Y	U	C	90
61	Y	O	C	30

Table B.3 continued

62	Y	U	C	90
63	Y	U	T	0
64	Y	O	T	-30
65	Y	D	C	90
66	Y	O	T	60
67	Y	O	C	90
68	Y	O	T	90
69	Y	O	C	30
70	Y	O	O	90
71	Y	O	O	-30
72	Y	D	C	30
73	Y	O	T	30
74	Y	O	C	90
75	Y	O	C	90
76	Y	O	C	0
77	Y	U	C	-60
78	Y	O	C	60
79	Y	U	T	30
80	Y	O	T	60
81	Y	D	C	30
82	Y	U	C	30
83	Y	O	T	-60
84	Y	O	C	-60
85	Y	O	T	-60
86	Y	O	C	90
87	Y	O	C	60
88	Y	U	C	90
89	Y	U	C	90
90	Y	D	C	30
91	Y	U	C	30
92	Y	O	C	-60
93	Y	U	C	-60
94	Y	O	C	30
95	Y	D	C	30
96	Y	O	T	90
97	Y	U	T	30
98	Y	U	C	30
99	Y	O	T	30
100	Y	O	T	30
101	Y	O	C	30
102	Y	D	C	-60
103	Y	U	C	-60
104	Y	D	C	30
105	Y	O	C	30
106	Y	D	C	90
107	Y	O	T	90
108	Y	U	C	0
109	Y	O	C	30
110	Y	U	C	90
111	Y	D	C	30
112	Y	D	C	30
113	Y	O	T	30
114	Y	D	C	90
115	Y	O	C	90
116	Y	O	T	90
117	Y	O	T	90
118	Y	O	C	90
119	Y	O	T	90
120	Y	O	T	0
121	Y	O	C	60
122	Y	D	C	90
123	Y	D	C	0

Data B.3 Statistics for Experiment with and without Influencer Strands

	Normalized Yield	Standard Error (Yield)	Normalized Hairpins|Rectangle	Standard Error (Hairpins)
No Influencers	0.9665715749	0.0196464811	0.9271559881	0.0374137523

	Staple Stacked (p Averaged)	Standard Error of p	Average Bernoulli SE	Scaffold Stacked (1-p Averaged)
No Influencers	0.3240225124	0.0165794166	0.019038137	0.6759774876
1 Staple Stacked	0.2183439213	0.021526932	0.0128769384	0.7816560787
2 Staple Stacked	0.2866341584	0.013455874	0.0161887529	0.7133658416
4 Staple Stacked	0.3947432171	0.0464785489	0.0238882708	0.6052567829
6 Staple Stacked	0.2565912576	0.0248912742	0.0148471129	0.7434087424
1 Scaffold Stacked	0.1456654016	0.0078853663	0.0102449419	0.8543345984
2 Scaffold Stacked	0.3009049741	0.0020317285	0.0195030601	0.6990950259
4 Scaffold Stacked	0.1857789518	0.0584629595	0.0116501337	0.8142210482
6 Scaffold Stacked	0.1783255429	0.0130800072	0.0115122921	0.8216744571
Average	0.2545566597			0.7454433403
Standard Error	0.0267107116			0.0267107116

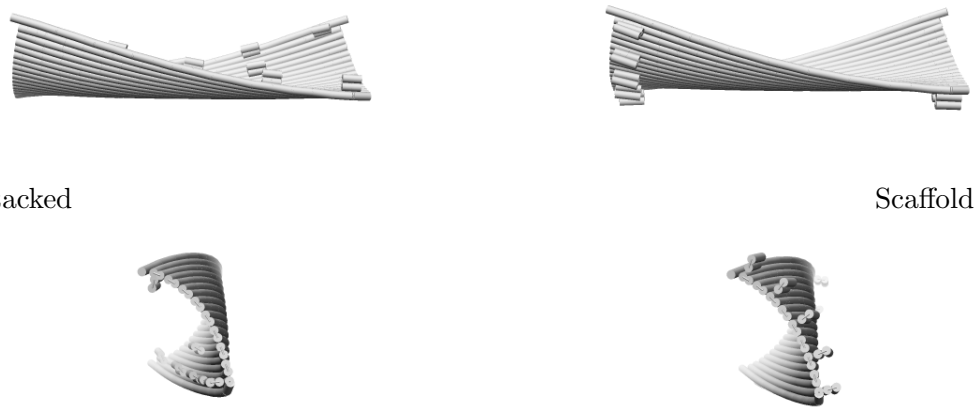
	Staple Stacked Up (Avg, norm)	Staple Stacked Down (Avg, norm)	Scaffold Stacked Up (Avg, norm)	Scaffold Stacked Down (Avg, norm)
No Influencers	0.903968254	0.096031746	0.2097365951	0.7902634049
1 Staple Stacked	0.9392712551	0.0607287449	0.4852408009	0.5147591991
2 Staple Stacked	0.7483660131	0.2516339869	0.3221597907	0.6778402093
4 Staple Stacked	0.8944444444	0.1055555556	0.5569023569	0.4430976431
6 Staple Stacked	0.7884615385	0.2115384615	0.6654693487	0.3345306513
1 Scaffold Stacked	0.8333333333	0.1666666667	0.3396621897	0.6603378103
2 Scaffold Stacked	0.7100815851	0.2899184149	0.2729640418	0.7270359582
4 Scaffold Stacked	NA	NA	NA	NA
6 Scaffold Stacked	0.8666666667	0.1333333333	0.4420493741	0.5579506259
Average	0.8355741363	0.1644258637	0.4117730622	0.5882269378
Standard Error	0.0284860748	0.0284860748	0.0542901721	0.0542901721

	St. Stk. -60 Avg, norm	St. Stk. -30 Avg, norm	St. Stk. 0 Avg, norm	St. Stk. 30 Avg, norm
No Influencers	0.0680076628	0.1791274033	0.2503710607	0.2963912188
1 Staple Stacked	0.0777216611	0.0303030303	0.2654320988	0.3080808081
2 Staple Stacked	0.1719452997	0.0714364345	0.1537035656	0.2789703381
4 Staple Stacked	0.1601525465	0.0761880413	0.0746434012	0.1196243861
6 Staple Stacked	0.0511378926	0.0251508861	0.0179045093	0.4065467266
1 Scaffold Stacked	0.1345373237	0.0371517028	0.3559339525	0.2041623667
2 Scaffold Stacked	0.201535585	0.0716805241	0.1045338213	0.0756189415
4 Scaffold Stacked	0.1096938776	0.0380527211	0.0829081633	0.487457483
6 Scaffold Stacked	0.0838815789	0.0471491228	0.1748903509	0.5422149123
Average	0.1176237142	0.0640266518	0.1644801026	0.3021185757
Standard Error	0.0173855437	0.0157409303	0.0361805314	0.0524017366

	St. Stk. 60 Avg, norm	St. Stk. 90 Avg, norm
No Influencers	0.0892875634	0.116815091
1 Staple Stacked	0.0976430976	0.2084736251
2 Staple Stacked	0.0885737926	0.2353705696
4 Staple Stacked	0.3993320763	0.1700595485
6 Staple Stacked	0.240427863	0.252421866
1 Scaffold Stacked	0.1012039904	0.1670106639
2 Scaffold Stacked	0.4240230205	0.1226081076
4 Scaffold Stacked	0.1785714286	0.1033163265
6 Scaffold Stacked	0.0663377193	0.0751096491
Average	0.187266728	0.1612428275
Standard Error	0.0461474382	0.0205450586

	Sc. Stk. -60 Avg, norm	Sc. Stk. -30 Avg, norm	Sc. Stk. 0 Avg, norm	Sc. Stk. 30 Avg, norm
No Influencers	0.0692460317	0.1521825397	0.1109126984	0.2875
1 Staple Stacked	0.0845524038	0.0552274067	0.1498921488	0.3702822513
2 Staple Stacked	0.1235577839	0.053310683	0.0653303719	0.391229054
4 Staple Stacked	0.2132359909	0.0759773729	0.0385493388	0.0509971142
6 Staple Stacked	0.1202097162	0.0420548341	0.0417335257	0.3714112554
1 Scaffold Stacked	0.029677113	0.0294396961	0.2141500475	0.3461538462
2 Scaffold Stacked	0.2617618761	0.2292564046	0.0605761199	0.0469391315
4 Scaffold Stacked	0.1261223926	0.0499240227	0.0420223788	0.4421605194
6 Scaffold Stacked	0.1081045752	0.0430065359	0.0882352941	0.4631372549
Average	0.1262742093	0.0811532773	0.0901557693	0.3077567141
Standard Error	0.0237514608	0.0220753967	0.0197883141	0.0517703862

	Sc. Stk. 60 Avg, norm	Sc. Stk. 90 Avg, norm
No Influencers	0.1109126984	0.2692460317
1 Staple Stacked	0.1440262843	0.1880790585
2 Staple Stacked	0.1230721584	0.2434999487
4 Staple Stacked	0.4578283155	0.1634118677
6 Staple Stacked	0.1897527658	0.2106435786
1 Scaffold Stacked	0.1709401709	0.2010921178
2 Scaffold Stacked	0.3161792538	0.0852872141
4 Scaffold Stacked	0.2679928167	0.0682690979
6 Scaffold Stacked	0.102875817	0.1780392157
Average	0.2092866979	0.1786186812
Standard Error	0.0392976499	0.0220990258



Staple Stacked Scaffold Stacked

Figure B.5 CanDo Tertiary Predictions

B.5 An Interesting Failure

For the peculiar correlation between being staple-stacked and facing up (in all experiments) or being scaffold-stacked and face down (in some experiments), we can explain the phenomenon by looking at the global twisting of the structure. Figure B.5 shows the predicted global twist for both the staple-stacked (on the left) and the scaffold-stacked (on the right). The correlations can be explained by a propensity for both conformations to bind with the mica on their convex side, as the convex side naturally has a greater surface area with which to bind.

An immediate consequence of global twisting arises from the dumbbell hairpins switching sides from one conformation to the other. The concavity might encourage the scaffold-stacked conformation through its dumbbells, even though we added the dumbbells as a reporting mechanism. Despite our intended use for dumbbells, their unintended influence may explain why the count of origami did not match our thermodynamic predictions.

One may then correctly assert, "if we could get rid of the global twist, we could also get rid of the unintended bias towards the scaffold-stacked structure." However, let's assume we added regular deletions in one of our CaDNAno designs to eliminate the global twist (Ke et al. (2009a)). Without loss of generality, let's assume we deleted every n^{th} nucleotide along each helix bundle from the scaffold-stacked conformation. Due to the winding pattern of the scaffold-stacked design, if we convert our structure to the staple-stacked design we will find that for

every n^{th} nucleotide we delete along all m helix bundles, we will have a helix bundle in the staple-stacked design that has m fewer nt than its neighbor bundles (due to the m rows of the rectangle causing m deletions along a winding segment of the scaffold). Such an origami is not energetically feasible. Thus simple methods for reducing global twist remove the multi-state nature of the origami. I leave this problem for future research.

BIBLIOGRAPHY

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, 92(3):291–294.

Amaldi, E., Iuliano, C., and Rizzi, R. (2010). Efficient deterministic algorithms for finding a minimum cycle basis in undirected graphs. In *Integer Programming and Combinatorial Optimization*, pages 397–410. Springer.

Andersen, E. S., Dong, M., Nielsen, M. M., Jahn, K., Lind-Thomsen, A., Mamdouh, W., Gothelf, K. V., Besenbacher, F., and Kjems, J. (2008). Dna origami design of dolphin-shaped structures with flexible tails. *ACS nano*, 2(6):1213–1218.

Andersen, E. S., Dong, M., Nielsen, M. M., Jahn, K., Subramani, R., Mamdouh, W., Golas, M. M., Sander, B., Stark, H., Oliveira, C. L., et al. (2009). Self-assembly of a nanoscale dna box with a controllable lid. *Nature*, 459(7243):73–76.

Bai, X.-c., Martin, T. G., Scheres, S. H., and Dietz, H. (2012). Cryo-em structure of a 3d dna-origami object. *Proceedings of the National Academy of Sciences*, 109(49):20012–20017.

Benson, E., Mohammed, A., Gardell, J., Masich, S., Czeizler, E., Orponen, P., and Högberg, B. (2015). Dna rendering of polyhedral meshes at the nanoscale. *Nature*, 523(7561):441–444.

Castro, C. E., Kilchherr, F., Kim, D.-N., Shiao, E. L., Wauer, T., Wortmann, P., Bathe, M., and Dietz, H. (2011). A primer to scaffolded dna origami. *Nature methods*, 8(3):221–229.

Chen, J. H., Churchill, M. E., Tullius, T. D., Kallenbach, N. R., and Seeman, N. C. (1988). Construction and analysis of monomobile dna junctions. *Biochemistry*, 27(16):6032–6038.

Devlin, K. (2003). *Sets, functions, and logic: an introduction to abstract mathematics*. CRC Press.

Dietz, H., Douglas, S. M., and Shih, W. M. (2009). Folding dna into twisted and curved nanoscale shapes. *Science*, 325(5941):725–730.

Douglas, S. M., Bachelet, I., and Church, G. M. (2012). A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, 335(6070):831–834.

Douglas, S. M., Dietz, H., Liedl, T., Högberg, B., Graf, F., and Shih, W. M. (2009a). Self-assembly of dna into nanoscale three-dimensional shapes. *Nature*, 459(7245):414–418.

Douglas, S. M., Marblestone, A. H., Teerapittayanon, S., Vazquez, A., Church, G. M., and Shih, W. M. (2009b). Rapid prototyping of 3d dna-origami shapes with cadnano. *Nucleic acids research*, page gkp436.

Doye, J. P., Ouldridge, T. E., Louis, A. A., Romano, F., Šulc, P., Matek, C., Snodin, B. E., Rovigatti, L., Schreck, J. S., Harrison, R. M., et al. (2013). Coarse-graining DNA for simulations of DNA nanotechnology. *Phys. Chem. Chem. Phys.*, 15(47):20395–20414.

Dunn, K. E., Dannenberg, F., Ouldridge, T. E., Kwiatkowska, M., Turberfield, A. J., and Bath, J. (2015). Guiding the folding pathway of dna origami. *Nature*.

Ellis-Monaghan, J. A., McDowell, A., Moffatt, I., and Pangborn, G. (2014). Dna origami and the complexity of eulerian circuits with turning costs. *Natural Computing*, pages 1–13.

Endo, M., Katsuda, Y., Hidaka, K., and Sugiyama, H. (2010). Regulation of dna methylation using different tensions of double strands constructed in a defined dna nanostructure. *Journal of the American Chemical Society*, 132(5):1592–1597.

Endo, M., Yamamoto, S., Emura, T., Hidaka, K., Morone, N., Heuser, J. E., and Sugiyama, H. (2014). Helical dna origami tubular structures with various sizes and arrangements. *Angewandte Chemie International Edition*, 53(29):7484–7490.

Fu, T. J. and Seeman, N. C. (1993). DNA double-crossover molecules. *Biochemistry*, 32(13):3211–3220.

Fuchs, J., Fiche, J.-B., Buhot, A., Calemczuk, R., and Livache, T. (2010). Salt concentration effects on equilibrium melting curves from dna microarrays. *Biophysical journal*, 99(6):1886–1895.

Geary, C. W. and Andersen, E. S. (2014). Design principles for single-stranded rna origami structures. In *DNA Computing and Molecular Programming*, pages 1–19. Springer.

Gerling, T., Wagenbauer, K. F., Neuner, A. M., and Dietz, H. (2015). Dynamic DNA devices and assemblies formed by shape-complementary, non–base pairing 3d components. *Science*, 347(6229):1446–1452.

Han, D., Jiang, S., Samanta, A., Liu, Y., and Yan, H. (2013). Unidirectional scaffold-strand arrangement in DNA origami. *Angewandte Chemie International Edition*, 52(34):9031–9034.

Huang, R. (2012). Rqda: R-based qualitative data analysis. r package version 0.2-3.

Jungmann, R., Avendaño, M. S., Woehrstein, J. B., Dai, M., Shih, W. M., and Yin, P. (2014). Multiplexed 3d cellular super-resolution imaging with dna-paint and exchange-paint. *Nature methods*, 11(3):313–318.

Ke, Y., Bellot, G., Voigt, N. V., Fradkov, E., and Shih, W. M. (2012). Two design strategies for enhancement of multilayer–dna-origami folding: underwinding for specific intercalator rescue and staple-break positioning. *Chemical Science*, 3(8):2587–2597.

Ke, Y., Douglas, S. M., Liu, M., Sharma, J., Cheng, A., Leung, A., Liu, Y., Shih, W. M., and Yan, H. (2009a). Multilayer dna origami packed on a square lattice. *Journal of the American Chemical Society*, 131(43):15903–15908.

Ke, Y., Sharma, J., Liu, M., Jahn, K., Liu, Y., and Yan, H. (2009b). Scaffolded dna origami of a dna tetrahedron molecular container. *Nano letters*, 9(6):2445–2447.

Kim, D.-N., Kilchherr, F., Dietz, H., and Bathe, M. (2012). Quantitative prediction of 3d solution shape and flexibility of nucleic acid nanostructures. *Nucleic acids research*, 40(7):2862–2868.

King, N. P., Sheffler, W., Sawaya, M. R., Vollmar, B. S., Sumida, J. P., André, I., Gonen, T., Yeates, T. O., and Baker, D. (2012). Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science*, 336(6085):1171–1174.

Kuzuya, A. and Komiyama, M. (2009). Design and construction of a box-shaped 3d-dna origami. *Chemical Communications*, (28):4182–4184.

Kuzuya, A. and Ohya, Y. (2014). Nanomechanical molecular devices made of dna origami. *Accounts of chemical research*.

Kuzuya, A., Sakai, Y., Yamazaki, T., Xu, Y., and Komiyama, M. (2011). Nanomechanical DNA origami 'single-molecule beacons' directly imaged by atomic force microscopy. *Nat Commun*, 2.

Lamsweerde, A. V. (2009). *Requirements Engineering - From System Goals to UML Models to Software Specifications*. Wiley.

Leveson, N. G. (1995). *Safeware: System Safety and Computers*. ACM, New York, NY, USA.

Levinthal, C. (1968). Are there pathways for protein folding. *J. Chim. phys*, 65(1):44–45.

Liedl, T., Högberg, B., Tytell, J., Ingber, D. E., and Shih, W. M. (2010). Self-assembly of three-dimensional prestressed tensegrity structures from dna. *Nature nanotechnology*, 5(7):520–524.

Linko, V. and Dietz, H. (2013). The enabled state of dna nanotechnology. *Current opinion in biotechnology*, 24(4):555–561.

Lund, K., Manzo, A. J., Dabby, N., Michelotti, N., Johnson-Buck, A., Nangreave, J., Taylor, S., Pei, R., Stojanovic, M. N., Walter, N. G., et al. (2010). Molecular robots guided by prescriptive landscapes. *Nature*, 465(7295):206–210.

Maňuch, J., Thachuk, C., Stacho, L., and Condon, A. (2011). Np-completeness of the energy barrier problem without pseudoknots and temporary arcs. *Natural Computing*, 10(1):391–405.

Marras, A. E., Zhou, L., Su, H.-J., and Castro, C. E. (2015). Programmable motion of DNA origami mechanisms. *Proceedings of the National Academy of Sciences*, 112(3):713–718.

Mathew-Fenn, R. S., Das, R., and Harbury, P. A. (2008). Remeasuring the double helix. *Science*, 322(5900):446–449.

Mathieson, L.-A. and Condon, A. (2015). On low energy barrier folding pathways for nucleic acid sequences. In *DNA Computing and Molecular Programming*, pages 181–193. Springer.

McNaught, A. D. and McNaught, A. D. (1997). *Compendium of chemical terminology*, volume 1669. Blackwell Science Oxford.

Nakayama, B. (2015). RQDA catalog of DNA origami failure diagrams. `http://www.cs.iastate.edu/~rlutz/cat`.

Nakayama, B. and Bahr, D. (2014). *Universal computation in the prisoners dilemma game*. Springer.

Petsko, G. A. and Ringe, D. (2004). *Protein structure and function*. New Science Press.

Peyret, N. (2000). *Prediction of Nucleic Acid Hybridization: Parameters and Algorithms*. PhD thesis.

Qian, L. and Winfree, E. (2011). Scaling up digital circuit computation with dna strand displacement cascades. *Science*, 332(6034):1196–1201.

Rothemund, P. W. (2006). Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302.

Saccà, B., Ishitsuka, Y., Meyer, R., Sprengel, A., Schöneweiß, E.-C., Nienhaus, G. U., and Niemeyer, C. M. (2015). Reversible reconfiguration of DNA origami nanochambers monitored by single-molecule fret. *Angewandte Chemie International Edition*, 54(12):3592–3597.

SantaLucia, J. (1998). A unified view of polymer, dumbbell, and oligonucleotide dna nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, 95(4):1460–1465.

SantaLucia Jr, J. and Hicks, D. (2004). The thermodynamics of dna structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415–440.

Schaeffer, J. M. (2012). *The multistrand simulator: Stochastic simulation of the kinetics of multiple interacting DNA strands*. PhD thesis, California Institute of Technology.

Scheible, M. B., Pardatscher, G., Kuzyk, A., and Simmel, F. C. (2014). Single molecule characterization of dna binding and strand displacement reactions on lithographic dna origami microarrays. *Nano letters*, 14(3):1627–1633.

Sinclair, J. F., Ziegler, M. M., and Baldwin, T. O. (1994). Kinetic partitioning during protein folding yields multiple native states. *Nature Structural & Molecular Biology*, 1(5):320–326.

Smith, S. B., Cui, Y., and Bustamante, C. (1996). Overstretching b-dna: the elastic response of individual double-stranded and single-stranded dna molecules. *Science*, 271(5250):795–799.

Sobczak, J.-P. J., Martin, T. G., Gerling, T., and Dietz, H. (2012). Rapid folding of DNA into nanoscale shapes at constant temperature. *Science*, 338(6113):1458–1461.

Stein, I. H., Schüller, V., Böhm, P., Tinnefeld, P., and Liedl, T. (2011). Single-molecule fret ruler based on rigid dna origami blocks. *ChemPhysChem*, 12(3):689–695.

Tun, T., Lutz, R., Nakayama, B., Yu, Y., Mathur, D., and Nuseibeh, B. (2015). The role of environmental assumptions in failures of dna nanosystems. In *Complex Faults and Failures in Large Software Systems (COUFLESS), 2015 IEEE/ACM 1st International Workshop on*, pages 27–33.

Wagenbauer, K. F., Wachauf, C. H., and Dietz, H. (2014). Quantifying quality in dna self-assembly. *Nature communications*, 5.

Walter, S. R., Young, K. L., Holland, J. G., Gieseck, R. L., Mirkin, C. A., and Geiger, F. M. (2013). Counting the number of magnesium ions bound to the surface-immobilized thymine oligonucleotides that comprise spherical nucleic acids. *Journal of the American Chemical Society*, 135(46):17339–17348.

Watson, J. D. and Crick, F. H. (1953). The structure of dna. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 18, pages 123–131. Cold Spring Harbor Laboratory Press.

Wickham, S. F., Bath, J., Katsuda, Y., Endo, M., Hidaka, K., Sugiyama, H., and Turberfield, A. J. (2012). A dna-based molecular motor that can navigate a network of tracks. *Nature nanotechnology*, 7(3):169–173.

Winfree, E. (1998). *Algorithmic self-assembly of DNA*. PhD thesis, California Institute of Technology.

Woo, S. and Rothemund, P. W. (2011). Programmable molecular recognition based on the geometry of dna nanostructures. *Nature chemistry*, 3(8):620–627.

Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M. D. (2006). Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic acids research*, 34(2):564–574.

Yoo, J. and Aksimentiev, A. (2013). In situ structure and dynamics of dna origami determined through molecular dynamics simulations. *Proceedings of the National Academy of Sciences*, 110(50):20099–20104.

Zhang, F., Jiang, S., Wu, S., Li, Y., Mao, C., Liu, Y., and Yan, H. (2015). Complex wireframe dna origami nanostructures with multi-arm junction vertices. *Nature nanotechnology*, 10(9):779–784.

Zhou, L., Marras, A. E., Su, H.-J., and Castro, C. E. (2015). Direct design of an energy landscape with bistable DNA origami mechanisms. *Nano letters*, 15(3):1815–1821.